tvst

Retina

Microperimetry Reliability Assessed From Fixation Performance

Amandeep Singh Josan[1,2], Isabella Farrance[1], Laura J. Taylor[1,2], Daniel Adeyoju[1,2], Thomas M. W. Buckley[2], Jasleen K. Jolly[1,3], and Robert E. MacLaren[1,2]

[1] Nuffield Laboratory of Ophthalmology, Nuffield Department of Clinical Neurosciences, University of Oxford, NIHR Oxford Biomedical Research Centre, Oxford, UK
[2] Oxford Eye Hospital, Oxford University Hospitals NHS Foundation Trust, Oxford, UK
[3] Vision and Eye Research Institute, Anglia Ruskin University, Cambridge, UK

Correspondence: Amandeep Singh Josan, Nuffield Laboratory of Ophthalmology, Level 6, John Radcliffe Hospital, West Wing, Headley Way, Headington, Oxford OX3 9DU, UK. email: amandeep.josan@ndcn.ox.ac.uk

Received: January 4, 2023
Accepted: April 23, 2023
Published: May 23, 2023

Keywords: microperimetry; reliability; BCEA; fixation

Citation: Josan AS, Farrance I, Taylor LJ, Adeyoju D, Buckley TMW, Jolly JK, MacLaren RE. Microperimetry reliability assessed from fixation performance. Transl Vis Sci Technol. 2023;12(5):21, https://doi.org/10.1167/tvst.12.5.21

Purpose: Microperimetry provides an accurate assessment of central retinal sensitivity due to its fundus-tracking capability, but it has limited reliability indicators. One method currently employed, fixation loss, samples the optic nerve blind spot for positive responses; however, it is unclear if these responses arise from unintentional button presses or from tracking failure leading to stimuli misplacement. We investigated the relationship between blind spot scotoma positive responses (termed *scotoma responses*) and fixation.

Methods: Part 1 of the study involved a custom grid of 181 points centered on the optic nerve that was constructed to map physiological blind spots in primary and simulated eccentric fixation positions. Scotoma responses and the 63% and 95% fixation bivariate contour ellipse areas (BCEA63 and BCEA95) were analyzed. In Part 2, fixation data from controls and patients with retinal diseases (234 eyes from 118 patients) were collected.

Results: Part 1, a linear mixed model of 32 control participants, demonstrated significant ($P < 0.001$) correlation between scotoma responses and BCEA95. In Part 2, the upper 95% confidence intervals for BCEA95 were 3.7 deg^2 for controls, 27.6 deg^2 for choroideremia, 23.1 deg^2 for typical rod–cone dystrophies, 21.4 deg^2 for Stargardt disease, and 111.3 deg^2 for age-related macular degeneration. Incorporating all pathology groups into an overall statistic resulted in an upper limit BCEA95 = 29.6 deg^2.

Conclusions: Microperimetry reliability is significantly correlated to fixation performance, and BCEA95 provides a surrogate marker for test accuracy. Examinations of healthy individuals and patients with retinal disease are deemed unreliable if BCEA95 > 4 deg^2 and BCEA95 > 30 deg^2, respectively.

Translational Relevance: Microperimetry reliability should be assessed using fixation performance as summarized by BCEA95 rather than the level of fixation losses.

Introduction

Microperimetry (also known as fundus-tracked perimetry) measures central retinal sensitivity and is now widely used as an outcome measure in clinical trials assessing novel treatments for inherited and acquired retinal disease.[1,2] Microperimetry combines the principles of standard static automated perimetry with an infrared scanning laser ophthalmoscope (SLO), alongside fundus-tracking software utilizing automatic landmark registration. The incorporation of an SLO allows retinal features to be viewed with an illumination source that is beyond the visible spectrum. This feature is paramount in allowing for simultaneous real-time fundus viewing without interfering with the stimulus presentation during testing. The result is dynamic stimuli placement that actively compensates for eye movements. Additionally, during blinks, head movements, or very large eye movements, the loss of landmark detection/registration and subsequent inability to place stimuli in the predefined location causes

the stimulus presentation to be automatically paused until fundus landmark registration is re-established. Microperimetry makes use of direct retinal stimuli using light-emitting diodes, whereas for conventional static perimetry stimuli are projected onto an external surface (cupola). These features cumulatively result in microperimetry being considered a more precise and consistent measure of retinal sensitivity compared to visual fields that are obtained via standard static perimetry. The ability to assess the retinal sensitivity of the same location on multiple visits to a greater degree of positional accuracy is an attractive prospect in clinical trials where assessment of localized longitudinal treatment effects are of great interest.

The fundus-tracking ability of both commonly available microperimeter devices, the MAIA (CenterVue, Padova, Italy) and the Nidek MP-1 and MP-3 microperimeters (Nidek Technologies, Padova, Italy), is limited by the 25-Hz (25 times per second) refresh rate of their respective SLO cameras. Although this is often assumed to be sufficient, it is important to consider that saccadic or microsaccadic eye movements have the potential to far exceed this tracking refresh rate. In addition, microsaccadic eye movements are likely to occur even in healthy individuals who are effectively fixating on the target.[3] A recent study on healthy individuals found that microsaccades occurred on average 0.84 times per second with a peak velocity of 43.68 deg/s.[4] With a 25-Hz SLO refresh rate, this would imply that, at these microsaccadic speeds, the eye could travel undetected by a theoretical 1.7° between fundus-tracking refresh events. A loss of tracking leading to a 1.7° margin of error in microperimetry grid placement would have significant implications in studies of diseases where there exists sharp demarcation between healthy and diseased retina such as in choroideremia. In addition, stimuli are presented for 200 ms, far exceeding the refresh rate of the fundus tracker and thus providing several opportunities for loss of tracking for the duration of the stimulus presentation. The situation may be particularly pertinent in eye diseases where fixation may be more compromised, such as in age-related macular degeneration or other maculopathies. In these conditions, larger amplitude saccades, tremors, or large fixational motions may be present, potentially increasing the tracking margin of error and significantly affecting the accuracy of stimuli placement and, hence, reliability of results.[5]

In the process of fundus tracking, the MAIA gathers a wealth of fixation data. The standard output provides a summary of the fixation performance during examination and displays several metrics to convey the results. Among these are the bivariate contour ellipse area (BCEA) encompassing 63% and 95% of all fixation points (BCEA63 and BCEA95, respectively). At present, however, microperimetry fixation data are rarely used in clinical trials, and their use is most commonly limited to investigational studies, such as changes in fixation with disease onset and progression,[6,7] or investigations into visual rehabilitation through fixation training in those who have lost foveal function.[8,9] To our knowledge, there have been no studies correlating BCEA fixation data to examination reliability, and there is no guidance on what level of fixation performance is acceptable. Data-gathering studies performed with normative BCEA fixation data collected in a multicenter study on a large cohort of 358 healthy volunteers have found an average BCEA95 of 2.40 degrees^2 $\pm$ 2.04 degrees^2.[10] However, the grid employed in the study was a 37-point radial grid pattern, a shorter duration test than that using the 68-point 10-2 regularly spaced grid pattern more commonly employed in clinical trials involving inherited retinal diseases.[11]

Although microperimetry has greatly advanced the assessment of central retinal sensitivity, current devices are not without limitations. During microperimetry examinations, there are no false-negative catch trials. The only measure of examination reliability given is via the fixation loss, metric which is evaluated via 10-dB stimuli presented to the center of the optic nerve head (Heijl–Krakau method[12]) at a rate of approximately once per minute. Although not explicitly stated, in light of the fundus-tracking abilities of microperimetry, these are assumed by most experts to be false-positive responses (or false-positive catch trials)—that is, due to accidental button presses in the absence of seen stimuli, implying an unreliable patient response.[10,13,14] Currently, microperimetry examinations are considered reliable if the number of fixation losses is below an upper limit of acceptance, often taken to be between 25% and 33%, depending on the protocol.[14,15] The MAIA operator's manual recommends a cut-off value of 30% for a test to be considered reliable. For a standard 10-2 examination typically lasting 8 to 10 minutes, this would represent approximately three button presses out of the eight to 10 stimuli presented to the optic nerve head.

It is important to acknowledge that these cut-off values have been arbitrarily adopted from glaucoma studies in static-automated perimetry where large databases of perimetry results were analyzed to establish the upper confidence limits of false-positive and false-negative catch trials.[16] These cut-off limits were selected to incorporate the largest proportion of examinations and minimize the number of rejected examinations. A false-positive and false-negative catch trial value of 30% to 33% was calculated as representing the upper 95% confidence intervals for the range of values encountered. This method prevented the exclusion of

excessively large numbers of normal and glaucoma examinations.[16–18]

Therefore, two important questions arise: (1) What does the fixation losses metric fundamentally represent? (2) Rather than applying a fixation losses cut-off value derived from glaucoma studies in static automated perimetry, what value constitutes an acceptable upper limit in the context of fundus-tracked microperimetry?

In this study, we propose that the optic nerve head responses given by the fixation losses metric, in addition to detecting false positives, may also have a contribution from undetected loss of fundus tracking, whereby stimuli are in fact seen due to the blind spot test being inadvertently presented outside of the optic nerve scotoma. If the source of these positive responses is significantly due to the latter explanation, then this would have far reaching implications in the presumption of accurate stimuli positional placement, not just during the blind spot test but for all stimuli presented throughout the microperimetry examination. In light of this latter possibility, we further hypothesize that fixation data collected throughout microperimetry testing, in the form of BCEA, may be a superior indicator of microperimetry reliability compared with the standard output fixation losses metric. To investigate the role of fixation performance on microperimetry reliability, we divided this study into two parts.

Part 1

In Part 1, we investigated whether the reported fixation losses metric on standard microperimetry output predominantly represents true false positives (unintentional button presses in the absence of seen stimuli) or undetected loss of fundus tracking (positive response resulting from misplacement of stimuli to a seeing region of the retina), which we term ***scotoma responses***. In the case of the latter, it would be reasonable to further hypothesize that the loss of fundus tracking causing scotoma responses is correlated to fixational eye movements. As such, one would expect that, as fixation performance declines, the number scotoma responses (misplaced stimuli) would increase in a predictable fashion. In order to thoroughly investigate this hypothesis, a custom grid was designed to significantly increase the number of stimuli placed at the blind spot. A range of healthy participants could then be examined under various situations designed to challenge fixation performance. Investigating fixational performance and its correlation to scotoma responses further, we also derived a new fixation metric related specifically to very large fixational motions which we term ***gaze spikes***. Our suggestion that the fundus-tracking abilities of the 25-Hz SLO may be a limiting factor would imply that those eye motions with the largest deviations from the mean may be closely correlated to the number of scotoma responses.

Part 2

Part 2 involved the construction of a database of microperimetry fixation data in the form of BCEA from (1) healthy control volunteers and (2) a range of ocular pathologies including choroideremia, retinitis pigmentosa (RP), Stargardt disease (STGD), and age-related macular degeneration (AMD), to inform the range of typical BCEA values encountered for each group. This database can then be used to assign criteria on acceptable BCEA values, which, in turn, infers the likely percentage of misplaced stimuli likely to arise given the typical fixational performance of each group. This can then be used to infer whether an examination for a particular pathology can be considered reliable specifically in the context of microperimetry. An accurate measure of microperimetry examination reliability is particularly pertinent for clinical trials where microperimetry may be used as a primary or secondary outcome measure to gauge novel treatment efficacy. Use of fixational data to assess microperimetry test reliability is an attractive prospect due to the continuous monitoring and extensive data collection on eye position that occur during an examination. This is contrasted with the infrequent blind spot presentation currently employed by the MAIA to assess reliability.

Methods

Part 1: Establishing Correlation Between Fixation and Reliability Using a Custom Testing Grid

Recruitment of healthy controls for Part 1 of this study was approved by the University of Oxford medical sciences interdivisional research ethics committee (ref. no. R77042/RE001) and adhered to the tenets of the Declaration of Helsinki. Participants below 18 years old or over 60 years were excluded, as well as any participants with pre-existing ocular conditions and spectacle prescriptions beyond the range of +4 to –8 diopters. Only the right eye was tested for each participant to minimize testing times and the compounding effects of fatigue.

Microperimetry testing was performed in a darkened room (light level < 1.0 lux), with no formal

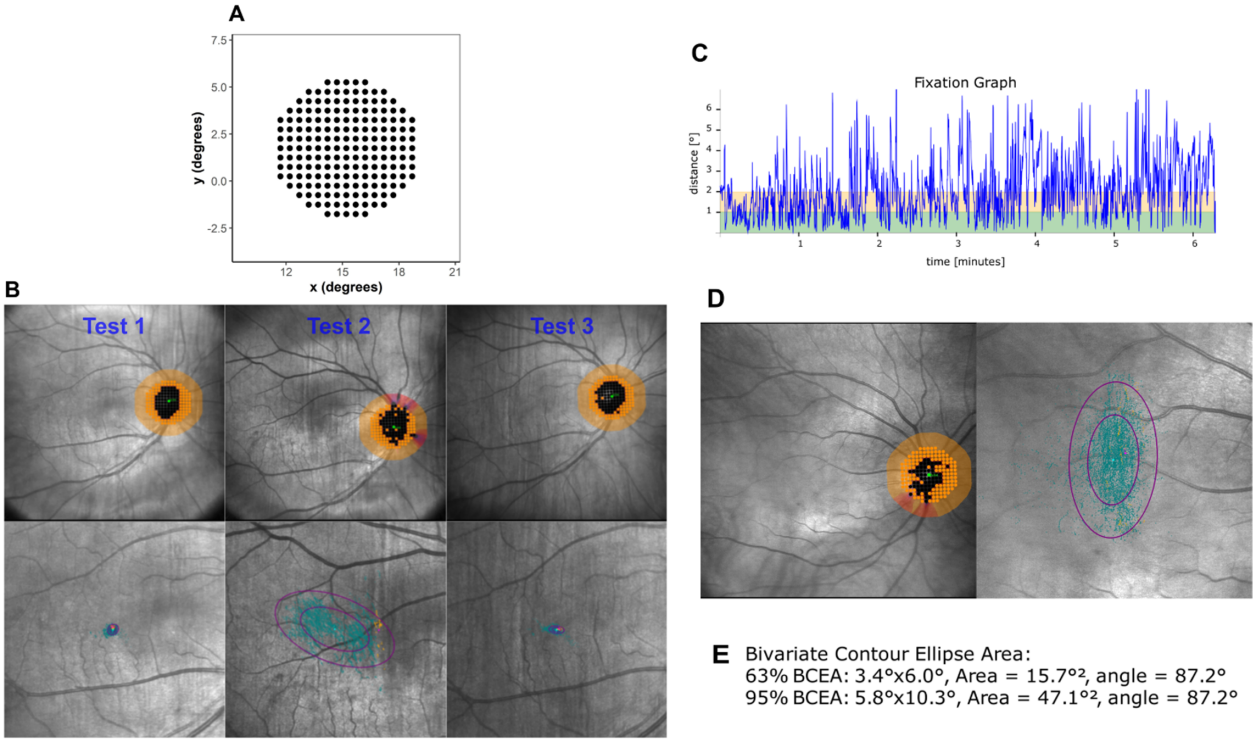

Figure 1. Healthy control example. (**A**) Custom 181-point grid pattern with 0.5° spacing. (**B**) Three consecutive examinations for an example participant: *left*, *middle*, and *right* with fixation on the target, fixation off of the target, and repeat fixation on the target, respectively. (**C**) Time-series fixation graph detailing eye position throughout testing. (**D**, *left*) Example participant result specifically highlighting the second (non-foveally centered) test only. (**D**, *right*) Corresponding eye positions throughout testing (*cyan dots*) with ellipse areas encompassing 63% and 95% of all points highlighted. (E) BCEA63 and BCEA95 areas and angle values.

dark adaption or pupil dilation. A novel custom testing grid made up of 181 points densely arranged (0.5° spacing) was used to map and approximately follow the optic nerve contours (Fig. 1A). This grid was constructed to be large enough so that approximately half of the test points fell outside of the optic nerve, eliciting a response, and approximately half were placed inside the optic nerve, theoretically eliciting no response. This ensured patient engagement throughout the test. The order of point testing was also sufficiently randomized to avoid long durations where no responses might occur. All stimuli were presented at the maximum brightness of the device (0 dB), as per the scotoma finder strategy, in order to effectively map the optic nerve scotoma.

Each participant completed three separate scotoma-finding strategy examinations (Fig. 1B) with a 5-minute break between each test. Microperimetry testing instructions were explained to each participant prior to examination, and verbal encouragement was provided throughout testing to ensure attentiveness. The first exam was initialized as a "new expert exam" using the 181-point custom grid manually centered on the participant's optic nerve. The participant was instructed to maintain fixation on the red fixation cross target, which was positioned on the temporal retina in order to bring the full extent of the optic nerve into view. The participant was instructed to press the button every time a flash of light was seen. The follow-up function was used for subsequent exams. For test two, the same instructions were repeated; however, in this instance, the participant was instructed to fixate eccentrically away from the target superiorly into a blank region of the screen. Care was taken to ensure that fixation was not directed so far as to displace the optic nerve from SLO view and preclude mapping. By discouraging the participant from fixating on any targets, the test would become significantly more difficult and the gaze would naturally tend to wander, resulting in a large spread of fixation. Each participant was warned prior to testing that this would be the case and were reassured. The third test repeated the protocol of the first test, with the participant focusing on the target again.

For each examination, data were collected detailing the seen and not seen stimuli (Fig. 1B), along with their locations, spread of fixation in the form of BCEA (Fig. 1E), and tracked time-series fixation graph

throughout testing (Fig. 1C). These data were exported as raw data export files (threshold.txt and fixation.txt) from the MAIA device.

Scotoma Responses

Of the three tests performed on each participant, the first was considered a learning test and not included in subsequent analyses. BCEA fixation data and scotoma responses from test 2 and test 3 were analyzed. The testing methodology assumes that eye movements that go undetected by the fixation tracker will be indicated through positive responses within the physiological scotoma of the optic disc as mapped by the custom grid. These scotoma responses are identical to the usual Heijl–Krakau blind spot fixation losses metric employed by the MAIA but at a far higher sampling rate. To differentiate these positive responses from the fixation losses metric employed by the MAIA, we refer to them as ***scotoma responses*** in order to specifically emphasize their unknown etiology.

To calculate the number of scotoma responses, it was paramount to identify the physiological center of the mapped optic nerve scotoma rather than the anatomical center of the optic nerve, as the two often did not coincide. The initial grid placement was carefully chosen to be at the center of the anatomical optic disc; however, subsequent results would often show a slight offset or shifting of the optic nerve scotoma from the anatomical center of the optic nerve. This may have been due to mild tilting of the optic discs or the presence of peripapillary atrophy, or it may have been due to minor errors in the manual placement of the custom grid at the initial setup phase. Indeed, it was not possible to align the testing grid such that a central positioning of the mapped scotoma could be guaranteed a priori. In fact, an offset scotoma was evident in almost all cases (see Fig. 2A). Hence, the center of the physiological optic nerve scotoma for each participant was identified manually (highlighted by a red star for the participant example shown in Fig. 2A). From this point, the surrounding 45 points were examined from the raw text files representing a region contained well within the scotoma for all participants. If any loci within these 45 points (Fig. 2B) of the center of the scotoma were logged as seen (assigned value of 1) rather than not seen (assigned value of 0), then these were defined as scotoma responses. The percentage of scotoma responses was calculated as a proportion of all 45 central points. Inclusion of only the central 45 points ensured that variability in responses from points displayed at the edge of the optic nerve scotoma (i.e., contentious responses arising from light bleed from scotoma border stimuli) did not contribute to the scotoma response results.

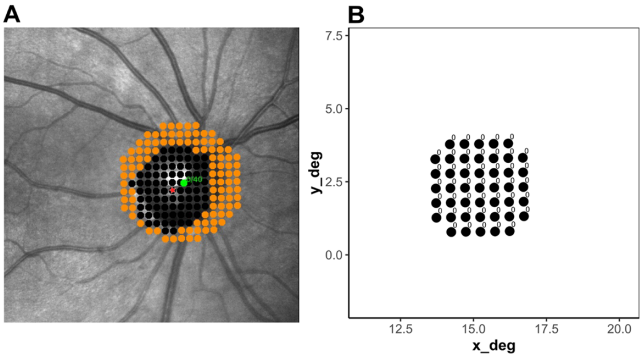

Figure 2. Healthy control example results. (**A**) A single right eye examination result demonstrating non-central position of optic nerve scotoma within a grid placed centrally on the anatomical optic nerve. In this case, the inferior temporal scotoma was likely due to the presence of a small band of peripapillary atrophy at the optic nerve border. The center of the scotoma was manually identified and is highlighted with a *red star*. (**B**) From this manually identified central scotoma point, the surrounding 45 points were isolated to assess the number of responses in this central region. These are defined as scotoma responses. In this case, there were zero scotoma responses.

Gaze Spikes

A new fixation metric termed ***gaze spikes*** was derived from the raw fixation data using the following procedure. The final preferential retinal locus (PRL_f), located automatically by the MAIA, is the average of all the PRL locations taken throughout the test. Using the PRL_f x and y locations as the reference points from which to subtract all prior fixation positions logged throughout the examination enables identification of any large movements, arbitrarily defined by deviations larger than 2°, at the time the stimulus was being presented. This is equivalent to counting the number of spikes above 2° as seen on the time-series fixation graph (Fig. 1C), with the additional criteria that we include only those that occurred during a stimulus presentation. We hypothesize that the number of recorded gaze spikes may be a surrogate marker for the number of undetected fixation deviations that would potentially lead to scotoma responses. Linear regression analysis with the number of gaze spikes as a dependent variable revealed non-normally distributed residuals. We therefore applied a generalized linear mixed model in this instance with a zero-inflated negative binomial distribution probability density function due to the propensity of zero gaze spike values. A zero-inflated model was compared with a conventional negative binomial and Poisson distribution functions, with the zero-inflated negative binomial outperforming on both residual and dispersion plots.

Zero-inflated fitted models have two components: (1) a zero-inflated component that assesses the probability of obtaining a zero versus non-zero value (logistic regression), and (2) a conditional component that performs a regression on those non-zero values. We report the conditional components of the model with the conditional R^2 calculated by comparing the full model with a null model. All of the relevant code used for this analysis is available for viewing on github (https://github.com/amanasj).

Data and Statistical Analysis

Analysis was performed using Prism 9.4.1 (GraphPad, San Diego, CA) and R 4.2.1 (R Foundation for Statistical Computing, Vienna, Austria),[19] as well as the lme4 and lmerTest packages.[20,21] A linear mixed model was used to model the relationship between BCEA fixation and the percentage of scotoma responses. BCEA63 and BCEA95 were set as the independent (predictor or explanatory) variables, and scotoma response values were set as the dependent (outcome) variable. In order to take into account the clustered nature of the data arising from duplicate testing of each participant, the participant ID was set as a random intercept variable. As such, the linear mixed model framework has the following format:

$$\begin{aligned} &\textit{Scotoma responses (per participant)} \\ &\quad = \textit{Slope} \times \text{BCEA} + \textit{Intercept (per participant)} \\ &\qquad + \textit{Residual error (per participant)} \end{aligned}$$

The linear regression model was used to analyze the normality of residuals in the dependent variable, and a generalized linear mixed model was used in instances where normality conditions were not met. A pseudo-correlation coefficient (pseudo-R^2) was established using a technique developed by Nakagawa and Schielzeth[22] which details the marginal and conditional correlation coefficients along with an overall coefficient value. P values were calculated using Satterthwaite-approximated degrees of freedom.[23]

Part 2: BECA Values in Different Patient Cohorts

To provide meaningful BCEA fixation criteria for assessing microperimetry reliability using a standard 10-2 grid, normative BCEA values for a range of participants were assessed as part of the Visual Function in Inherited Retinal Disease study (International Standard Randomised Controlled Trial Number 24016133), approved by the UK Health Regulatory Authority (ref. no. 20/WM/0283). The right and left eyes for each patient were included in the analysis by again using a linear mixed model with each participant set as the random intercept variable. However, in this instance, the independent variable was omitted, allowing for a model in which the intercept represents the overall mean of each group while controlling for clustering arising from the within-subject nature, reflecting the likely correlated measures between right and left eyes per participant.

Results

Part 1: Scotoma Responses and Fixation Performance

A total of 32 healthy participants were recruited and underwent triplicate microperimetry examination using the 181-point optic nerve mapping custom grid. Six participants were excluded, as they were not able to complete the microperimetry testing due to small pupils. Fifty-two exams from 26 participants were used in subsequent analyses. Figure 3A shows a scatterplot of BCEA95 versus scotoma responses obtained from test 2 and test 3. The colored lines represent the relationship between the independent predictor variable (BCEA95) and the dependent outcome variable (scotoma responses) for a selection of participants taken at random to illustrate the mixed modeling process. The displayed subsets of individual regression lines have intercepts that are free to vary for each participant, accounting for the nested nature of repeat testing and individual variability in examination results on a participant level. The solid-black fitted line represents the overall regression (fixed effects only) and is representative of the mean regression line across the whole cohort. The equation of line is given by this black line but with variances derived from the range of individual random intercept lines.

The independent predictor variables, BCEA63 and BCEA95, were found to be almost perfectly correlated with each other ($R^2 = 1$, $P < 0.001$) (Fig. 3B), in agreement with the findings of Morales et al.,[10] justifying the use of BCEA95 alone in subsequent analyses for simplicity. We caution that this near perfect correlation suggests that the BCEA63 and BCEA95 ellipses may have deliberately been fitted by MAIA with common major and minor axes and angles rather than being fitted independently of each other. The results of the fixation metrics, BCEA95 and gaze spikes, as independent predictor variables plotted against the scotoma response percentages are shown in Figure 4. BCEA95 was found to be significantly correlated to the percentage of scotoma responses (BCEA95: pseudo $R^2 = 0.64$, $P < 0.001$) (Fig. 4A). The number of gaze spikes

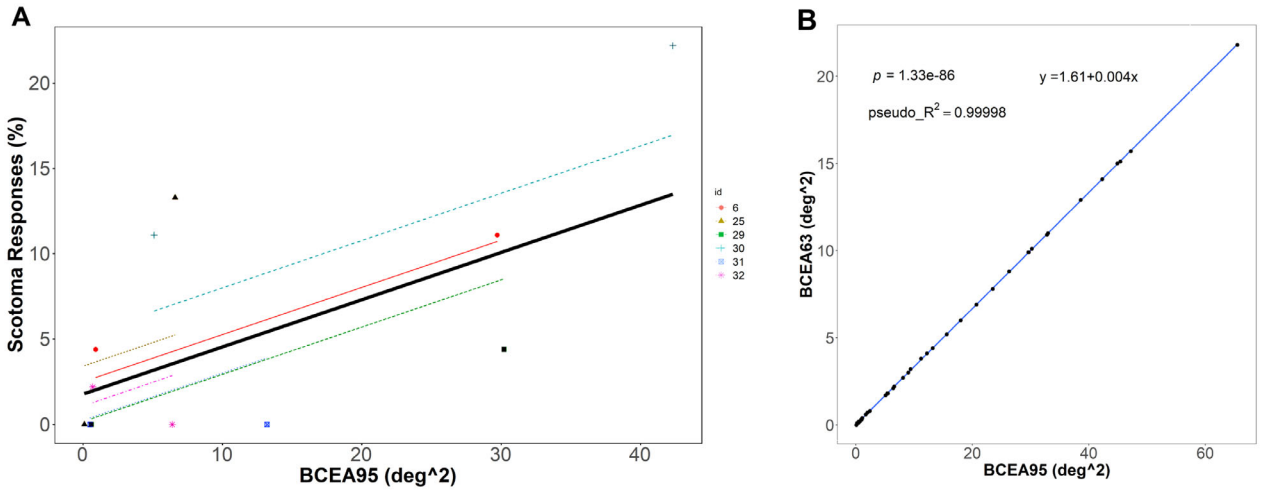

Figure 3. Linear mixed-model visualizations. (**A**) Visualization of the linear mixed regression model for a selection of participants, fitted with random intercepts as indicated by the *colored lines*. *The black line* represents the overall cohort mean regression line. (**B**) Linear mixed-model regression analysis demonstrating an almost perfect correlation between BCEA63 and BCEA95.

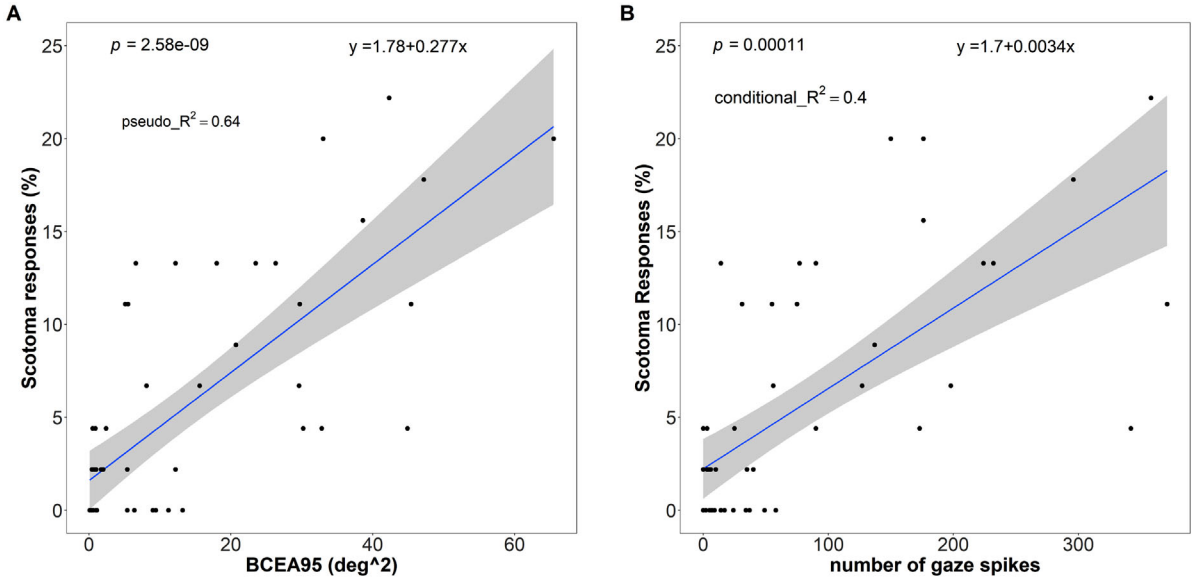

Figure 4. Linear mixed-model regression model for the scotoma response percentage dependent outcome variable versus the independent predictor variables. (**A**) BCEA95. (**B**) Number of gaze spikes. The *shaded regions* represent the standard errors. *P* values, pseudo R^2, and conditional R^2 correlation coefficients and the equations of lines are displayed alongside each plot.

was also significantly correlated to the percentage of scotoma responses ($P < 0.001$, conditional $R^2 = 0.4$) (Fig. 4B).

With further analysis of the linear regression model with BCEA95 as the predictor (Fig. 4A), the intercept implied that, for BCEA95 = 0 (i.e., hypothetically perfect fixation), an average of 1.78% (SD = 2.25) scotoma responses were still predicted to remain. As the fundus tracker should operate perfectly in this fictional scenario, the 1.78% average scotoma responses may therefore be cautiously interpreted to be the average number of unintentional button presses across the whole cohort. This corresponds to a very reasonable three out of 181 unintentional button presses (true false positives in the presence of a correctly placed unseen stimulus). Applying the MAIA-recommended cut-off value of 30% to the fixation losses in the equation of line shown in Figure 4A would suggest that this value of scotoma responses (equivalent to a constantly sampled version of the fixation loss index) translates to a BCEA95 value of 102 deg^2.

Part 2: Range of BCEA and Inferred Microperimetry Examination Reliability

A total of 234 microperimetry examinations from 118 eyes were analyzed across five different groups of participants/patients: control (n = 77 eyes), choroideremia (n = 68 eyes), RP (n = 44 eyes), STGD (n = 27 eyes), and AMD (n = 18 eyes). From these, the distribution of BCEA95 values for each group are shown using violin plots (Fig. 5). The mean and 95% confidence intervals for each group are given in the Table along with combined statistics for the right eyes and left eyes.

As seen in the Table, BCEA95 values for the control group were extremely low with a very narrow range of values, as would be expected for a cohort with healthy foveal function. The mean BCEA95 (± SD) for the control group was 2.7 ± 3.3 deg^2, which is in good agreement with findings in previous studies using radial grid patterns.[10,24,25] Hence, assuming a reliable examination is one where a given examination demonstrates a BCEA95 value that falls within the upper 95% confidence intervals for the control group, our findings indicate that the BCEA reliability threshold for the control group is BCEA95 = 3.7 deg^2. For illustration purposes, using Figure 4A, we can infer that this upper

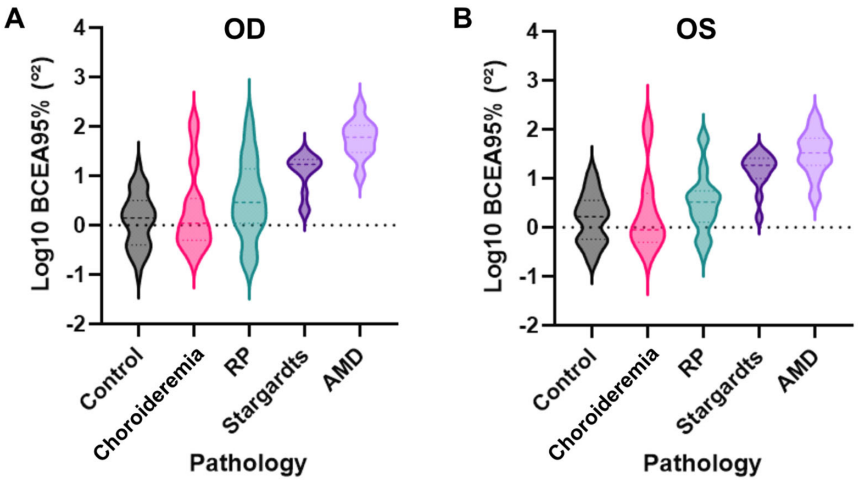

Figure 5. Grouped fixation performance plots. Violin plots display the log BCEA95 values for the right eye (**A**) and left eye (**B**) for the control and pathology groups. The fovea of the STGD and AMD groups have a larger spread of fixation (larger BCEAs) than the more parafoveal/peripheral-affected choroideremia and RP groups.

Table. Grouped Descriptive Statistics

	Control		Choroideremia		RP		STGD		AMD	
	OD	OS	OD	OS	OD	OS	OD	OS	OD	OS
Group statistics										
Sample size (n)	39	38	34	34	22	22	13	14	9	9
BCEA95, mean (SD)	2.5 (3.2)	2.8 (3.5)	14.8 (34.8)	17.5 (41.2)	15.5 (32.8)	9.4 (18.2)	15.3 (8.5)	17.8 (10.0)	81.8 (76.0)	57.5 (47.5)
Lower 95% CI	1.5	1.7	2.7	3.1	0.9	1.3	10.2	12.0	23.4	21.0
Upper 95% CI	3.6	4.0	27.0	31.9	30.0	17.5	20.4	23.6	140.2	94.0
Combined (OD and OS) repeated-measures statistics										
Combined mean (deg^2)	2.7		16.2		12.4		16.5		69.7	
Combined lower 95% CI	1.7		4.7		1.8		11.6		28.0	
Combined upper 95% CI	3.7		27.6		23.1		21.4		111.3	
Scotoma response (%) (inferred misplaced stimuli)	2.8		9.4		8.2		7.7		32.6	
OD vs. OS mean difference P value (repeated-measures analysis of variance)	0.41		0.67		0.12		0.27		0.09	

The top section displays individual group statistics for OD and OS, and the bottom section displays combined OD and OS statistics obtained by evaluating a linear mixed model to account for the within-subject repeated measurements. Units of mean BCEA95 and confidence intervals (CIs) are deg^2. The inferred percentages of misplaced stimuli (scotoma responses) were evaluated from the equations of line shown in Figure 4.

limit of reliability corresponds to a predicted scotoma response value (or percentage of misplaced stimuli) of 2.8%. The consequence of this is that, for controls, the fundus tracker is generally very capable at actively compensating for the small gaze deviations encountered in this group and is able to accurately place stimuli in the intended locations. However, if a control participant demonstrated a BCEA95 > 4 deg^2, then that person would fall outside the expected 95% confidence intervals for this group and so the examination should be considered unreliable. Further, from Figure 4A we can predict that, for an examination with BCEA95 = 4 deg^2, this level of fixation instability would lead to total number of misplaced stimuli of around 2.8% of all presented stimuli.

Similarly, for the ocular pathology groups, the BCEA95 upper 95% confidence limits for the choroideremia, RP, STGD, and AMD groups were 27.6 deg^2, 23.1 deg^2, 21.4 deg^2, and 111.3 deg^2, respectively. As we can see, the AMD group had a range of BCEA values with an upper 95% confidence limit with an inferred number of misplaced stimuli given by the scotoma responses at 32.6%. This is close to the 30% fixation losses value recommended by the MAIA manufacturer that has been taken from glaucoma studies and is often used in many current microperimetry studies and clinical trials. All other values, however, are considerably smaller, with the implication that current guidelines are likely far too lenient, resulting in the inclusion of a large number of unreliable tests in many studies.

No significant differences among the groups were found in the mean OD and OS BCEA95 values, implying symmetrical fixation performance, likely as a consequence of symmetry in disease progression between each eye. This further adds to evidence of the contralateral eye as a suitable control in studies involving microperimetry. Combining all ocular pathology groups (OD and OS) into a single group statistic (i.e., combining all groups minus controls) using a mixed modeling process revealed an overall mean BCEA95 = 21.3 deg^2 (95% confidence interval, 13.1–29.6).

Discussion

It is commonly assumed that the fundus-tracking ability incorporated within microperimetry has largely eliminated the issue of eye movements during examination. In this study, we have designed a custom grid to vastly increase the sampling of the optic nerve scotoma compared to the infrequent sampling via the Heijl–Krakau method employed in the 'fixation losses' metric used in microperimetry. We have provided convincing evidence that increasing fixation instability causes gradual and predictable increases in scotoma responses, suggesting a failure of eye movement detection by the fundus tracker. Although the increased examination difficulty with simulated eccentric fixation (test 2) may also lead to more unintentional button presses, it is highly unlikely that the number of these would increase in a predictable and linear in fashion, as fixation worsens and BCEA95 increases. We, therefore, excluded false-positive button presses (due to trigger-happy participants) as being the significant cause of these blind-spot scotoma responses in favour of misplacement due to loss of fundus tracking as the primary cause.

Unwanted eye movements are a familiar problem that has confounded the interpretation of results from static automated perimetry since its inception. The accuracy and repeatability of retinal sensitivity measurements have vastly improved with the adoption of microperimetry devices; however, as we have shown, fixation stability is still an important variable in reliability and highlights the limitations of the fundus-tracking feature, particularly in patients who are unable to maintain stable fixation. The strong correlation between fixation performance and examination reliability (evaluated through scotoma responses) adds significant evidence that large and rapid saccadic (or microsaccadic) eye movements during microperimetry lead to an increased probability of undetected losses of fundus tracking, thereby bypassing the protocol of the device to pause the exam. This would lead to undetected stimuli misplacement and high pointwise test–retest variability[15] and may explain the increased variability seen at the borders of deep scotomas.[26] Although this issue has the potential to affect all patients undergoing microperimetry, it is particularly those with poor fixation who are more likely to have undetected misplaced stimuli, leading to a higher percentage of misplaced stimuli and hence increasingly unreliable results. These findings emphasize the importance of encouraging patients to maintain their gaze as steady as possible throughout testing, rather than assuming that the fundus tracker will sufficiently compensate for gaze deviations.

Current reliability indices and criteria in microperimetry (namely, the fixation losses metric) have been adopted from large glaucoma studies using static automated perimetry. It is, however, highly doubtful that there is any validity in using these cut-off values in microperimetry studies. Primarily, and arguably most significantly, the addition of a fundus tracker dramatically increases the placement accuracy of stimuli to predefined locations when compared with standard perimetry devices. Second, the standard 10-2 grid assesses only retinal sensitivity within the central

20° of retina. Static automated perimetry, in contrast, usually examines a much larger region with inclusion of the peripheral visual field. Due to the difficulty in testing increasingly peripheral locations, variability would be expected to increase at larger eccentricities, and overall variability may reasonably be expected to be higher in glaucoma studies. This would be reflected in more relaxed reliability cut-off criteria being applied in static automated perimetry. Adopting the same reliability cut-off criteria in microperimetry would therefore result in excessive numbers of unreliable examinations being deemed acceptable, when, in fact, stricter criteria should be applied to reflect the greater accuracy of the microperimetry device.

Establishing the correlation between scotoma responses and BCEA95, we expected to see a strengthening of the correlative relationship between scotoma responses and the newly derived metric, number of gaze spikes. However, the results of our analysis showed this not to be the case. Although both metrics were significant in their correlation to scotoma responses, it is perhaps BCEA95 that is a more representative measure of overall fixation performance, with the result being that BCEA95 is a better predictor of those specific fixation deviations that lead to grid placement errors resulting in scotoma responses.

From Part 2 of this study, taking results across the various ocular pathology groups studied, it is evident and intuitive that precise BCEA95 cut-off criteria for reliability assessment can be categorized based on the presence or absence of foveal function; however, in practice, it may be difficult to assess foveal function prior to microperimetry testing. Additionally, in early choroideremia and RP, normal foveal function is often observed, with fixation only affected late in the disease phase (as evident in the spread of BCEA95, shown in Fig. 5). We therefore propose that guidance on reliability based on BCEA95 should be categorized into two broad groups: (1) healthy controls (2) those with retinal disease. For the purposes of reliability assessment, the pathology groups can be viewed as a single group with varying levels of retinal disease and, as such, a diverse range of BCEA95 fixation performance. Conversely, the healthy control group should have very stable fixation across all demographics (e.g., age, ethnicity, gender), thereby eliminating possible confounding effects. As such, a much tighter bound on BCEA95 can be consistently applied to this group. Hence, for healthy controls, a test should be deemed reliable only if BCEA95 < 4 deg^2 (conservatively rounded up from 3.7 deg^2), and those with any retinal disease should be deemed reliable only if BCEA95 < 30 deg^2 (conservatively rounded up from 29.6 deg^2). If BCEA95 values surpass these cut-off values, then the results should be deemed unreliable despite the quoted number of fixation losses.

The study limitations include the following: (1) A relatively small database was used to derive the BCEA cut-off values in each disease group in order to evaluate test reliability criteria. These may not be comprehensively representative of the phenotype or severity that may be encountered for each disease group or for the derived overall statistics. (2) A limitation of the regression line analysis (Fig. 4) is that it is only strictly valid for the range of values shown. Extrapolating beyond these ranges may be invalid if the correlation alters at extreme values of BCEA not encountered in our study, although we consider this unlikely. (3) One final potential limitation is the hypothetical scenario where a patient may exhibit large-amplitude but slow-speed gaze motions away from the target throughout testing. In this case, a large BCEA95 would be recorded, but the fundus tracker would likely be able to consistently monitor and maintain tracking. Hence, in this instance, the microperimeter would likely be capable of accurate grid positioning despite a large recorded BCEA95 value. In this hypothetical scenario, the correlation between BCEA and scotoma responses would break down and lead to an erroneous assumption of poor reliability using our criteria. In practice, we have not encountered this situation, but the possibility should be noted, particularly in children where attentiveness may play a role in fixation performance.

We propose that fixation loss is a poor overall reliability index due to the extremely infrequent sampling of the optic nerve scotoma. High fixation losses in those with intact foveal fixation and a low BCEA95 value would likely represent true false positives, thus detecting those patients who may be excessively trigger happy. However, high fixation losses in those with poor fixation and high BCEA95 values are more likely to represent misplacement of stimuli due to a loss of fundus tracking. We therefore propose that fixation performance in the form of BCEA95 should be used primarily as an overall measure of examination quality and reliability with fixation losses, at best, considered a combined measure of fundus-tracking loss and false-positive catch trials. Both indices, forming part of the standard MAIA output, are readily available and so can be easily adopted by a quick inspection after an examination has been completed.

We emphasize that caution should be exercised in examining patients with nystagmus, as it is likely that the incorporated fundus tracker would be unable to compensate sufficiently and will likely lead to many misplaced stimuli and poor reliability and repeatability. Furthermore, the use of microperimetry to

monitor patients with AMD may have to be cautiously reviewed and use restricted to only those who are able to maintain a reasonably steady gaze (regardless of whether they are able to see the fixation target). In such cases, it may be prudent to perform several training tests to ensure that the patient is able to consistently reach the reliability threshold of BCEA95 < 30 deg^2 in the presence of eccentric fixation.

Conclusions

Assessing the reliability of microperimetry examinations is of paramount importance in clinical trial settings where the efficacy of potential new treatments is determined by changes in retinal sensitivity. Evidence-based reliability criteria in microperimetry are currently lacking, and the sources of errors are poorly understood. The often-used 30% fixation losses cut-off criterion employed to distinguish a reliable test from one that is unreliable has been adopted from investigations using static automated perimetry, which differs significantly from microperimetry. This study has demonstrated that, despite the significant improvement in measures of retinal sensitivity by the use of fundus-tracking technology, limitations exist that can be quantified by the fixation performance of a patient.

This study proposes the use of BCEA95 rather than fixation losses as the metric of choice to assess microperimetry examination quality and reliability due to the insufficient sampling rate of the latter index. In control patients, a reliability cut-off criterion of BCEA95 < 4 deg^2 is recommended. In patients with any form of retinal diseases with the potential to significantly affect foveal function (such as choroideremia, RP, STGD, or AMD), one could enforce a less stringent fixation cut-off criterion reflecting the spectrum of possible disease and fixation states; hence, a test would be deemed reliable if BCEA95 < 30 deg^2. This less stringent BCEA95 < 30 deg^2 can be applied to all groups (including controls) to simplify and keep consistent the percentage of acceptable misplaced stimuli for all. It should be noted, however, that as controls would be expected to perform considerably better than the ocular pathology groups on average, a more lenient cut-off for controls could have the consequence of the inclusion of potentially unreliable tests. Conversely, the more demanding BCEA95 < 4 deg^2 cut-off could be applied to all patient groups; however, it should be expected that a greater number of patients with foveal disease would fail this criterion, leading to the unintended consequence of excluding many late-stage potential patients from clinical trial enrollment.

Acknowledgments

The authors thank Pete Jones for helpful discussions during the analysis of data.

Supported by the National Institute for Health and Care Research (NIHR) Oxford Biomedical Centre and by a grant from the NIHR under its Research for Patient Benefit (RfPB) Programme (NIHR202821 to LT). The views expressed are those of the authors and not necessarily those of the National Health Service, the NIHR, or the Department of Health and Social Care. The sponsor and funding organization had no role in the design or conduct of this research.

Disclosure: **A.S. Josan**, None; **I. Farrance**, None; **L.J. Taylor**, None; **D. Adeyoju**, None; **T.M.W. Buckley**, None; **J.K. Jolly**, None; **R.E. MacLaren**, None

References

1. Buckley TMW, Jolly JK, Josan AS, Wood LJ, Cehajic-Kapetanovic J, MacLaren RE. Clinical applications of microperimetry in RPGR-related retinitis pigmentosa: a review. *Acta Ophthalmol*. 2021;99(8):819–825.
2. Pfau M, Jolly JK, Wu Z, et al. Fundus-controlled perimetry (microperimetry): application as outcome measure in clinical trials. *Prog Retin Eye Res*. 2021;82:100907.
3. Rucci M, Poletti M. Control and functions of fixational eye movements. *Ann Rev Vis Sci*. 2015; 1:499–518.
4. Montes S, Bennett D, Bensinger E, et al. Characterizing fixational eye motion variance over time as recorded by the tracking scanning laser ophthalmoscope. *Transl Vis Sci Technol*. 2022;11(2):35.
5. Davies A, Nanda A, MacLaren RE. Effect of reduced visual acuity on microperimetry performance in patients with inherited retinal degenerations. *Invest Ophthalmol Vis Sci*. 2019;60(9):1820–1820.
6. Grenga PL, Fragiotta S, Cutini A, Meduri A, Vingolo EM. Microperimetric evaluation in patients with adult-onset foveomacular vitelliform dystrophy. *Indian J Ophthalmol*. 2017;65(5):385–389.
7. Ozen Tunay Z, Idil A. Macular sensitivity assessment and fixation analysis using microperimetry in

children with retinopathy of prematurity. *Br J Ophthalmol*. 2022;106(12):1767–1771.

8. Ramírez Estudillo JA, León Higuera MI, Rojas Juárez S, de Lourdes Ordaz Vera M, Santana YB, Suazo BC. Visual rehabilitation via microperimetry in patients with geographic atrophy: a pilot study. *Int J Retina Vitreous*. 2017;3:21.
9. Ratra D, Gopalakrishnan S, Dalan D, Ratra V, Damkondwar D, Laxmi G. Visual rehabilitation using microperimetric acoustic biofeedback training in individuals with central scotoma. *Clin Exp Optom*. 2019;102(2):172–179.
10. Morales MU, Saker S, Wilde C, et al. Reference clinical database for fixation stability metrics in normal subjects measured with the MAIA microperimeter. *Transl Vis Sci Technol*. 2016;5(6):6.
11. Cehajic-Kapetanovic J, Xue K, Martinez-Fernandez de la Camara C, et al. Initial results from a first-in-human gene therapy trial on X-linked retinitis pigmentosa caused by mutations in RPGR. *Nat Med*. 2020;26(3):354–359.
12. Heijl A, Krakau CE. An automatic perimeter for glaucoma visual field screening and control. Construction and clinical cases. *Albrecht Von Graefes Arch Klin Exp Ophthalmol*. 1975;197(1):13–23.
13. Montesano G, Naska TK, Higgins BE, Wright DM, Hogg RE, Crabb DP. Determinants of test variability in scotopic microperimetry: effects of dark adaptation and test indices. *Transl Vis Sci Technol*. 2021;10(1):26.
14. Yang Y, Dunbar H. Clinical perspectives and trends: microperimetry as a trial endpoint in retinal disease. *Ophthalmologica*. 2021;244(5):418–450.
15. Wu Z, Ayton LN, Guymer RH, Luu CD. Intrasession test–retest variability of microperimetry in age-related macular degeneration. *Invest Ophthalmol Vis Sci*. 2013;54(12):7378–7385.
16. Yohannan J, Wang J, Brown J, et al. Evidence-based criteria for assessment of visual field reliability. *Ophthalmology*. 2017;124(11):1612–1620.
17. Heijl A, Patella VM, Bengtsson B. *The Field Analyzer Primer: Effective Perimetry*, 4th ed. Dublin, CA: Carl Zeiss Meditec; 2012.
18. Vingrys AJ, Demirel S. False-response monitoring during automated perimetry. *Optom Vis Sci*. 1998;75(7):513–517.
19. R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing; 2020.
20. Bates D, Mächler M, Bolker B, et al. Fitting linear mixed-effects models using lme4. *J Stat Softw*. 2015;67(1):1–48.
21. Kuznetsova A, Brockhoff PB, Christensen RHB. lmerTest package: tests in linear mixed effects models. *J Stat Softw*. 2017;82(13):1–26.
22. Nakagawa S, Schielzeth H. A general and simple method for obtaining R2 from generalized linear mixed-effects models. *Methods Ecol Evol*. 2013;4(2):133–142.
23. Luke SG. Evaluating significance in linear mixed-effects models in R. *Behav Res Methods*. 2017;49(4):1494–1502.
24. Molina-Martín A, Piñero DP, Pérez-Cambrodí RJ. Reliability and intersession agreement of microperimetric and fixation measurements obtained with a new microperimeter in normal eyes. *Curr Eye Res*. 2016;41(3):400–409.
25. Montesano G, Rossetti LM, Allegrini D, Romano MR, Garway-Heath DF, Crabb DP. Systematic and random mapping errors in structure - function analysis of the macula. *Transl Vis Sci Technol*. 2021;10(2):21.
26. Wu Z, Jung CJ, Ayton LN, et al. Test–retest repeatability of microperimetry at the border of deep scotomas. *Invest Ophthalmol Vis Sci*. 2015;56(4):2606–2611.