

Repeatability of Microperimetry in areas of RPE and Photoreceptor loss in Geographic Atrophy supported by AI-based OCT biomarker quantification

Leonard M. Coulibaly , Klaudia Birner , Azin Zarghami , Markus Gumpinger , Simon Schürer-Waldheim , Philipp Fuchs , Hrvoje Bogunović , Ursula Schmidt-Erfurth , Gregor S. Reiter

PII: S0002-9394(24)00520-8
DOI: <https://doi.org/10.1016/j.ajo.2024.11.005>
Reference: AJOPHT 13138

To appear in: *American Journal of Ophthalmology*

Received date: February 12, 2024
Revised date: September 27, 2024
Accepted date: November 6, 2024

Please cite this article as: Leonard M. Coulibaly , Klaudia Birner , Azin Zarghami , Markus Gumpinger , Simon Schürer-Waldheim , Philipp Fuchs , Hrvoje Bogunović , Ursula Schmidt-Erfurth , Gregor S. Reiter , Repeatability of Microperimetry in areas of RPE and Photoreceptor loss in Geographic Atrophy supported by AI-based OCT biomarker quantification, *American Journal of Ophthalmology* (2024), doi: <https://doi.org/10.1016/j.ajo.2024.11.005>



Short Title: Repeatability of Microperimetry in GA based on OCT-Biomarker

Repeatability of Microperimetry in areas of RPE and Photoreceptor loss in Geographic Atrophy supported by AI-based OCT biomarker quantification

*Leonard M. Coulibaly¹, Klaudia Birner¹, Azin Zarghami¹, Markus Gumpinger², Simon Schürer-Waldheim², Philipp Fuchs¹, Hrvoje Bogunović², Ursula Schmidt-Erfurth^{*1,2}, Gregor S. Reiter¹*

Surname indicated by underline."

- 1) Department of Ophthalmology and Optometry, Medical University of Vienna, Vienna, Austria
- 2) Laboratory for Ophthalmic Image Analysis, Department of Ophthalmology and Optometry, Medical University of Vienna, Vienna, Austria

Keywords: Fundus controlled perimetry; Microperimetry; Test-retest repeatability; Automated OCT-Biomarker quantification; Geographic atrophy; AMD; Photoreceptor integrity

Funding: Research presented in this manuscript received no funding.

Corresponding Author marked by * (also address for reprints):

Prof. Ursula Schmidt-Erfurth, MD
Department of Ophthalmology and Optometry
Medical University of Vienna
Währinger Gürtel 18-20
1090 Vienna, Austria
ursula.schmidt-erfurth@meduniwien.ac.at

Tel: +43 01 40400 79310

Fax: +43 01 40400 79320

Wordcount: 4321

5 Figures and 4 Tables 1 Supplement-File

Abstract

Purpose: Growing interest in microperimetry (MP) or fundus-controlled perimetry (FCP) as targeted psychometric testing method in geographic atrophy (GA) is warranted due to the disease subclinical/extra-foveal appearance or preexisting foveal loss with visual acuity becoming unreliable. We provide comprehensive pointwise test-retest repeatability reference values on the most widely used MP devices and combine them with targeted testing in areas of retinal pigment epithelium (RPE) as well as photoreceptor (PR) integrity loss, guiding the interpretation of sensitivity loss during the long-term follow-up of GA patients.

Design: Prospective reliability study

Methods: Patients with GA underwent consecutive testing on CenterVue (iCare) MAIA and NIDEK MP3 devices. Obtained PWS measurements were spatially co-registered to an optical coherence tomography (OCT) volume scan acquired during the same visit. Areas with RPE and PR integrity loss, drusen and PR thickness as well as the volume of hyperreflective foci (HRF) were identified and quantified using a set of validated deep learning-based algorithms. Test-retest repeatability was assessed according to areas defined by biomarker-specific morphologic changes using Bland-Altman coefficients of repeatability (CoR). Furthermore, the inter-device correlation, the repeatability of scotoma point detection as well as any potential effects on fixation stability were assessed.

Results: 900 stimuli per device from twenty subjects were included. Identical overall PWS test-retest variance could be detected for MAIA (± 6.57) and MP3 (± 6.59). PR integrity loss was associated with a higher test-retest variance on both devices (MAIA: $p=0.002$; MP3: $p<0.001$). Higher CoR for stimuli in areas presenting RPE loss (± 10.99 vs ± 5.34) or HRF (± 9.21 vs ± 6.25) could only be detected on MP3 examinations ($p<0.001$ and $p=0.01$, respectively). An excellent intra-device correlation (MAIA: 0.94[0.93-0.95] MP3: 0.94[0.94-0.95]) and a good mean inter-device correlation (0.84[0.53-0.92]) could be demonstrated. The chosen device, run order or absence of foveal sparing had no significant effect on fixation stability.

Conclusion: Areas presenting automatically quantified PR integrity loss with and without underlying RPE loss are associated with higher test-retest variance for both MAIA and MP3. These findings are crucial for an accurate interpretation of GA progression during long-term follow-up and the planning of future trials with microperimetry testing as functional study endpoint.

Background

The clinical appearance of geographic atrophy (GA), or non-neovascular atrophy secondary to age-related macular degeneration (AMD), is marked by clearly demarcated areas with degenerative changes of the photoreceptor (PR) layers and the underlying retinal pigment epithelium (RPE) as observed on histopathologic assessments¹ and optical coherence tomography (OCT) imaging.² This progressive neurodegenerative process has a multi-faceted pathophysiology, but eventually leads to an irreversible loss of function in affected retinal areas.³ With many millions of reported cases worldwide⁴, continuous aging of the general population and limited therapeutic options to slow disease progression⁵, GA must be considered one of the leading causes of retinal blindness.⁶ Regarding the recent FDA approval of complement inhibitors reducing the growth rate of atrophic lesions secondary to AMD⁷, the need for reliable functional endpoints to evaluate therapy effectiveness has become a focus of attention.⁸ Best corrected visual acuity (BCVA) assessment using the Early Treatment Diabetic Retinopathy Study (ETDRS) Chart remains the most common functional outcome in clinical trials examining therapy effectiveness. The strength of this method lies in its widespread accessibility and high practicality. However, VA testing has challenging limitations, as it reflects only the functionality of the central macular area. Given the frequent extrafoveal appearance of atrophic lesions in early GA, it becomes evident that VA alone is an insufficient clinical study endpoint to assess disease progression.⁹ Moreover, a wipe-out of the fovea of any extension would always result in unchanged BCVA values over time despite progression. Regulatory agencies have emphasized the need for additional functional clinical trial endpoints to complement VA assessment.¹⁰

Microperimetry (MP) also called fundus-controlled perimetry (FCP) is a well-established psychometric testing method, assessing retinal function across the entire macular region. By combining SLO-fundus imaging with integrated eye-tracking software and pointwise light stimulation, it allows for comprehensive and customizable retinal sensitivity mapping. Multiple manufacturers have developed and commercialized MP devices, offering mesopic, photopic, or scotopic testing conditions depending on the device and manufacturer.¹¹ Recent publications have highlighted the viability of MP in detecting subclinical changes in early disease stages of both intermediate and late-stage AMD.¹² Given its frequent non-foveal manifestation in the early stages and loss of fixation ability in advanced disease stages, the applicability of MP in the functional assessment of GA is evident.⁹ Considering the large number of potentially treatable patients, the identification of OCT imaging biomarkers

predicting accelerated GA lesion growth has become a major area of research.^{13,14} Photoreceptor (PR) integrity loss has been described to precede and exceed atrophic lesion growth and therefore constitutes one of the most promising predictive biomarkers for disease progression.^{1,15-17} The FDA considers the assessment of PR, as the functional correlate in morphology, and the loss of PR as an approvable clinical endpoint assessing GA progression, further laying emphasis on the importance of quantitative photoreceptor monitoring.¹⁸ The presence of drusen¹⁹ and hyperreflective foci (HRF)²⁰ have been listed as potential indicators for an accelerated disease progression.

Recently, multiple studies have proposed MP examinations as functional outcome in regulatory clinical trials examining potential therapeutic options for GA.²¹ However, comprehensive reference values for the test-retest repeatability of MP examinations in subjects with GA, using two commercially widely available MP devices, have not been published on the same study cohort. Most importantly, novel algorithms identifying PR integrity loss on the OCT image have not been considered systematically. The study presented here aims to fill this knowledge gap by providing test-retest repeatability values with regards to underlying retinal morphology, for a correct interpretation of disease progression in routine clinical settings and for the design of future clinical trials. This analysis also explores an inter-device comparison, evaluating the interchangeability of MP devices from different manufacturers under their standard settings.

Methods

This prospective, cross-sectional study was executed under the tenets of the declaration of Helsinki and was approved by the independent ethics committee of the Medical University Vienna. Patients with diagnosed GA secondary to AMD, were eligible for the inclusion in this analysis. Therefore, all included patients presented at least a complete RPE and outer retina atrophy (cRORA) lesion with 250 μm of continuous RPE loss and underlying hyper-transmission, along with overlying PR degeneration on OCT imaging. Both GA locations, patients with and without foveal sparing, were included into the study. Exclusion criteria comprised the presence of exudative signs on OCT, previous anti-VEGF treatments, any media opacity, advanced cataract, any signs of glaucoma (c/d ratio > 0.7 or a history of ocular high pressure) or concomitant maculopathies in the study eye. Only one eye per patient could be included. In cases where both eyes were eligible, the one with the better OCT imaging quality was selected. All participating patients were recruited from the out-patient clinic for retinal disease of the department of ophthalmology at the University Hospital Vienna and provided written informed and provided written informed consent before any study-related procedures were performed.

Study Procedures

Study eligibility was firstly determined by reviewing the patients' medical history and if no exclusion criteria were met, an additional assessment of a $20\times 20^\circ$ volumetric scan obtained from a spectral-domain OCT, the Heidelberg Spectralis HRA+OCT device (Heidelberg Engineering, Heidelberg, Germany) was conducted. After inclusion into the study, mydriatic eye drops (0.5% Tropicamid) were administered, to ensure proper pupil dilatation. Subsequently, two consecutive MP examinations on both the MP-3 (NIDEK CO., Ltd., Gamagori, Japan) and MAIA (MAIA, CenterVue (iCare) S.p.A., Padova, Italy) on the study eye (in total four MP examinations). A randomization tool determined which device was used first. A block size of 4 guaranteed similar number of eyes in both groups. To prevent fatigue, a mandatory break of at least ten minutes was held between each examination. All examinations took place on the same day, in a dark, windowless room with illumination levels below 1 lux, an eye patch on the fellow eye and were performed by a single experienced examiner.

Both devices employed the same automated stimulation pattern including 45 stimuli centred around the fovea and a 4-2 staircase strategy (see Figure 1). The stimulus size was set to Goldmann III, lasting for 200 milliseconds. The baseline MP3 examinations started with a

stimulus set at 17 decibels (dB) in each quadrant. The second examination on each device used the integrated follow-up function. To ensure consistency with a typical clinical setting, the standard testing mode was selected for each device. It is worth noting that the CenterVue (iCare) MAIA exclusively offered mesopic testing with a background of 4 asb (1.27 cd/m²) and a maximal luminance of 1000 asb (318.3 cd/m²), while the MP-3 employed photopic testing with a background of 31.4 asb (10 cd/m²) and a maximal luminance of 10000 (3183 cd/m²) asb under its standard setting.

Figure 1. 45 stimuli study grid for MP examinations

Automated imaging Biomarker analysis and registration

OCT Biomarker analysis:

Disease-specific imaging biomarkers were automatically segmented by previously published and validated deep learning algorithms.^{17,22–25}

The used algorithm for detecting RPE loss at an A-scan level is based on a convolutional neural network (CNN) featuring projective skip connections, which served to condense and transform encoded 3-dimensional OCT features into a 2-dimensional en-face binary map that in the case of Spectralis device can be put in direct spatial correspondence with the SLO image (see Figure 2: D).²² Applying the same technical principle, a discontinuation of the EZ layer between top of the ellipsoid zone and the outer boundary of the interdigitation zone defined as PR integrity loss was detected (see Figure 2: E). A separate CNN-based layer segmentation algorithm using the same boundaries calculated a PR-thickness map (see Figure 2: F) over the complete macular region encompassing each B-scan.²⁶ Similarly, a drusen thickness map (see Figure 2: C) was computed based on the layer segmentation of the outer boundary of the retinal pigment epithelium (OBRPE) and the Bruch's membrane (BM).²⁷ Finally HRF were defined as dot-shaped lesions with a volume larger than 0.06 nl with equal or higher reflectivity than the RPE and were segmented within the complete neurosensory retina above the RPE with a dedicated CNN (see Figure 2: B).^{28,29}

Figure 2. Automated imaging biomarker segmentation

Co-registration between Microperimetry and OCT volume:

An in-house developed algorithm was used to perform the point-to-point co-registration of each MP stimuli to the corresponding OCT volume scan. The pixelwise positioning of each stimulus was detected for the MAIA and MP3 on the SLO image and colour fundus photography, respectively. A specifically developed image registration algorithm, which

matched prominent retinal structures, spatially registered each stimulus point on the corresponding SLO image of the OCT volume scan. The vessels segmentation is based on the detection of the vessel junctions is achieved with Mask R-CNN based on the published work of Arikian et al.^{30,31} Figure 3 presents an illustration of the co-registration of microperimetry results on OCT imaging. All registrations were visually inspected and if necessary, manually corrected by marking matched prominent retinal structures and determining the transformation matrix with the ‘least squares’ method, by a single rater, with extensive knowledge on multimodal image registration. All pixels and their related imaging biomarkers within a radius of 70 μm were considered around each stimuli point.

Figure 3: Co-registration of microperimetry stimuli

Statistical analysis

Test-retest repeatability was assessed using Bland-Altman plots and corresponding coefficients of repeatability (CoR) at a 95% probability for stimuli with and without RPE loss, PR integrity loss, presence of drusen thickness and HRF volumes. Further, all measurements were divided between 1) areas with RPE and PR loss, 2) areas with PR loss and no RPE loss and 3) areas without RPE or PR integrity loss. The effect of PR and drusen thickness on test-retest repeatability was assessed by dividing the observed measurements into decreasing thickness quartiles, 0-25% (I.Q), 25-50% (II.Q), 50-75% (III.Q), 75-100% (IV.Q) of values. For the calculation of drusen thickness quartiles, the areas without drusen (Drusen thickness = 0) were omitted from the analysis. As Bland-Altman coefficients of repeatability were calculated by multiplying the standard deviation of the within subject stimuli-point difference by 1.96³², a Levene-test assessing homogeneity of variances was used to determine significant differences between subgroups encompassing stimuli with presence or absence of quantified biomarkers. Paired t-tests were used to calculate differences between mean retinal sensitivity, mean reaction time and average testing time between examinations.

Separate Bland-Altman plots were used to investigate repeatability of scotoma point detection for both devices after applying a 0 dB, 5 dB and 10 dB threshold.

The inter-device correlation was analysed using a linear regression model as well as interclass correlation coefficients (ICC) estimates and their 95% confidence intervals based on a mean-rating ($k = 2$), 2-way mixed-effects model with absolute-agreement for all examinations.³³ Linear mixed models with run (first or second performed examination for each device), chosen device and presence of foveal sparing as fixed factors estimated any potential effects on fixation stability.

All statistical analyses were performed using SPSS Statistics and only p-values below 0.05 were deemed significant.

Results

900 stimuli per device from twenty eyes of twenty subjects with a mean age of 77 (± 6) years were included in this analysis. RPE loss, i.e. clinical atrophic lesion size ranged from 0.46 to 7.12 mm² with a mean of 2.30 ± 1.86 mm². As expected, the PR integrity loss exceeded the RPE loss ranging from 1.49 to 13.09 mm² with a mean of 4.87 ± 3.35 mm². 30% (6/20 eyes) of included patients presented no foveal sparring, defined as an involvement of the foveal center point in the atrophic lesion. The MAIA was used first in 8 patients, while the MP3 was used first in 12 patients. Repeatability was almost identical between first and second used device (CoR: ± 6.56 dB vs ± 6.58 dB, respectively).

Mean retinal sensitivity (SD) for the first MAIA and MP3 examination was 18.05 (3.84) dB and 21.41 (3.28) dB respectively. For the follow up examination it was 18.35 (3.38) dB for MAIA and 21.79 (3.09) dB for MP3. No significant difference between first examination and follow up could be detected for both devices (MAIA: $p=0.76$; MP3: $p=0.72$).

Average testing and reaction time

Average testing times for MAIA examinations were 417 (40) and 370 (28) seconds for the baseline and follow-up examination respectively. For MP3 the baseline examination took an average of 506 (187) seconds and the follow-up examination an average of 422 (107) seconds. Mean reaction time (SD) were almost identical between baseline and follow-up MAIA examination (528 (68) and 532 (62) milliseconds, respectively with no significant difference ($p=0.34$).

Intra-device repeatability

Repeatability metrics from the Bland-Altman plots for each biomarker and the respective number of stimuli located within the area are summarized under Table 1. Fifteen stimuli points for examinations on MAIA and one stimulus point for MP3 could not be matched as they lay outside the OCT volume scan and were therefore omitted from the analysis.

*Table 1 Test-retest repeatability according to geographic atrophy imaging biomarker; Subgroups with heterogeneous variances are marked by an **

Impact of atrophy location: The presence or absence of foveal sparing had no significant effect on repeatability (MAIA: $p=0.67$; MP3: $p=0.3$).

Impact of RPE and/or PR integrity loss: Whether the pointwise retinal sensitivity measurement was taken within an area of RPE and simultaneous PR integrity loss, within a zone with solely PR integrity loss or an area without any sign of RPE or PR integrity loss had a significant impact (both $p<0.001$) on repeatability for MAIA and MP3. In accordance with this, a significant drop in repeatability with higher CoR can be observed between retinal areas with and without PR integrity loss for both devices ($p=0.002$ for MAIA and $p<0.001$ for MP3). Meanwhile, presence of RPE led to a significant increase of retest variance for measurements taken on MP3 ($p<0.001$) but not on MAIA ($p=0.3$).

Impact of HRF and drusen presence: Similarly to RPE loss, a significant increase of retest variance for HRF presence measurements could be detected on repeated measurements on MP3 ($p<0.001$) but not on MAIA ($p=0.66$). For both devices the presence or absence of drusen (MAIA: $p=0.09$; MP3: $p=0.35$) had no significant effect on repeatability.

Impact of PR thickness: A significant difference between PR layer thickness quartiles could be detected for all quartiles in MP3 ($p<0.001$) and for the thinnest 3 quartiles in MAIA ($p=0.014$). The detailed results for the PR thickness analysis can be found in Table 2.

Impact of drusen thickness: The thickness of drusen had no significant effect on repeatability as the Levene-test detected homogeneous variances between all quartiles (MAIA: $p=0.24$; MP3: $p=0.51$). Repeatability according to drusen thickness are summarized under Table 2.

Table 2. Test-retest repeatability according to photoreceptor and drusen thickness quartiles, Subgroups with heterogeneous variances are marked by an *.

Scotoma detection repeatability

Generally, an absolute scotoma with loss of retinal sensitivity is defined by the incapability of the subject to recognize the brightest stimulus. In case of MAIA and MP3 the floor of physical dynamic range is defined as -1 dB and 0 dB respectively. Meanwhile recent publications have reported an effective dynamic range floor of 10 dB for PWS assessment on MAIA.³⁴

Repeatability of scotoma detection between repeated testing was assessed using Bland-Altman plots for ≤ 0 dB (Figure 4: A and B), ≤ 5 dB (Figure 4: C and D), and ≤ 10 dB (Figure 4: E and F) on both devices. All Bland-Altman plots showed similar limits of agreement

Figure 4. Bland-Altman plots for repeatability of scotoma detection

Intra- and Inter-device correlation

ICCs [95% CI] between each of the devices are summarized in Table 3. Significant correlations (all $p < 0.05$) could be found between every performed examination. While the intra-device correlation can be considered excellent (MAIA: 0.937 [0.928-0.945] MP3: 0.944 [0.936-0.951]), only a good mean inter-device correlation could be detected (0.833 [0.528-0.917]).

Table 3. Interclass correlation coefficients between each performed microperimetry examination

The inter-device correlation using a linear regression model ($R^2=0.715$) can be summarized under the following regression equation:

$$\text{Mean MAIA} = 0.79 * \text{Mean MP3} + 1.126$$

Figure 5 is a graphic representation of the linear regression model between mean PWS from first and second examination results for MP-3 and MAIA examinations and therefore consists of 900 measurements.

Figure 5. Graphical representation of linear regression model

Fixation stability

The median fixation stability (IQR) for MAIA was 87.3 % (12.1) within the central 2° (P1) and 98.05 % (4.3) within the central 4° (P2). MP3 examinations had a median fixation of 90% (10) within the central 2° and 98.5 % (4.8) within the central 4°. The chosen device or whether it was the first or second examination on each device, termed Run 1 (first examination) and Run 2 (second examination), had no significant impact on fixation stability. Further presence of foveal sparing had no significant effect on fixation stability. Results are summarized under Table 4. As median fixation stability was similar between MAIA and MP3, an average of both devices was used during the assessment of the impact of acquisition sequence (Run1/Run2) and presence/absence of foveal sparing.

Table 4. Results for fixation stability in central 2° and 4° according to device, first or second examination (Run), presence or absence of foveal sparing

Journal Pre-proof

Discussion

The growing interest in MP as a targeted functional assessment method for GA is justified, given its unique capability to provide a topographic retinal sensitivity mapping over the complete macular region relatable to morphological changes at the level of the RPE and PR layers. Pivotal prospective clinical trials, evaluating the therapeutic effect of disease slowing drugs have considered changes in MP's sensitivity assessment as important secondary study outcome.^{35,36} This study provides much needed comprehensive reference values for mesopic as well as photopic MP device repeatability in patients with GA.

An almost identical PWS repeatability could be detected for MAIA (CoR: ± 6.57) and MP3 (CoR: ± 6.59) examinations over the complete macular region. Compared to healthy individuals we observed a drop in repeatability as a previously published analysis by our group detected a CoR of ± 4.61 and ± 4.55 in healthy patients over for MAIA and MP3 respectively.³⁷ Alibhai et al. published a lower PWS CoR of ± 3.58 dB for MAIA in GA performed on the same day. Notably, the used study grid included more equally spaced stimuli (93 stimuli) compared to our study grid (45 stimuli).³⁸ A higher number of stimuli points might therefore encompass more "healthy" retinal tissue and therefore lead to less variance in repeated measurements. Further some degree of our detected higher CoR may be related to repeated testing on multiple devices during the same day inducing potential patient weariness. Pfau et. al. published a test-retest repeatability assessment for examinations on MAIA, using a patient tailored study grid and reported a similar CoR of ± 6.64 dB and lower retest variance in areas with higher sensitivity.³⁹ In our analysis we detected comparable CoR results as well as a similar interaction between lower test-retest variance and higher sensitivity (see Table 1).

Previous publications examining the use of MP in the monitoring of GA have laid emphasis on the border-zone of the GA lesions. The area surrounding the atrophic lesion is often described as outer junctional zone^{40,41} or perilesional area⁴² and typically presents varying degrees of degenerative alterations of retinal cell layer morphology.⁴³ There is no unanimous definition of the extent of this border-zone, but it is often delineated as within 500 μ m of the atrophy border.⁴⁴

However, defining a junctional zone using a metric distance surrounding the atrophic lesion, as in the above-mentioned example, fails to account for the high interpatient variability. Traditionally, the definition of atrophic lesions was based on histology¹ and an FAF-based

GA border definition², in the light of advanced definitions of GA lesions by AI-based image analysis on OCT, we defined the junctional zone surrounding the atrophic lesion as objectively quantified PR integrity loss exceeding the margins of the RPE loss on OCT imaging. OCT imaging offers valuable 3D information on neurosensory morphology in contrast to the 2D information provided by FAF. This distinct junctional zone of atrophy, which in GA consistently originates from primary PR loss, is of utmost importance as this area indicates future GA progression.²³

Areas presenting a combination of PR integrity and RPE loss (within the GA lesion) as well as areas with solely PR integrity loss (often surrounding the GA lesion) had a significantly poorer test-retest repeatability compared to areas without these two GA specific morphologic alterations. This effect could be observed on examinations performed on MAIA as well as MP3. The combination of personalized AI-based biomarker quantification and targeted MP measurements in the individual areas affected by PR integrity and/or RPE loss will be crucial to further our understanding of the potential protective effects that novel therapeutics could have on the junctional zone around the atrophic lesion. Most important, establishing references for test-retest variance of PWS in the personalized areas of PR integrity and RPE loss, compared to a metric absolute general definition of the junctional zone, allows for a more precise and individualized understanding of longitudinal data in clinical trials.

Interestingly, a large difference between areas with PR integrity loss alone (CoR: ± 6.24) and areas with a combination of PR+RPE loss (CoR: ± 10.98) could be detected on MP3 with an almost doubling of the retest variance. Similarly, the presence of RPE loss alone had a notable impact on re-test variance only in MP3 testing. Neither of these two findings could be derived from examinations performed with the MAIA. We hypothesize that this difference might be attributed to the photopic testing conditions under which MP3 examinations were held.

Although counterintuitive to the assumption of an absolute scotoma, detecting a certain limited degree of retinal sensitivity within the atrophic lesion (area with RPE and PRI loss), as we did in this analysis, is in accordance with other analyses performed in the literature. Pfau et al. demonstrated that light sensitivity might be detected in a so called inner junctional zone defined as a patient-tailored iso-hull of -0.645° within the atrophic lesion.⁴⁰ Histological analysis has demonstrated that cone photoreceptor nuclei can be detected within the atrophic lesion in the form of outer retinal tubulations (ORT) or as a remaining part of the outer nuclear layer (ONL).⁴⁴

Photopic more than mesopic testing condition might stimulate these remaining cone photoreceptors nuclei and lead to an irregular detection of retinal sensitivity within affected areas. Additionally, stray light effects might be more pronounced with higher luminance and increase retest variance. Our findings regarding the repeatability of absolute scotoma detection underline these hypotheses, as a non-negligible degree of re-test variance in scotoma detection can be observed no matter the scotoma threshold definition (≤ 0 , ≤ 5 or ≤ 10) on both devices.

The correlation between PR thinning and an increase of CoR was further emphasized in our sub-analysis of PR-thickness quartiles. A clear trend of increased test-retest variance with decreasing levels of PR-thickness can be observed. However, it is worth noting, that the thickest PR-layer quartile on MAIA examination lay outside that trend. A previously published analysis on inter-device repeatability in healthy subjects, attributed a significantly higher retest variance to the central millimeter compared to other macular areas on MAIA examinations.³⁷ In the here presented analysis we did not calculate retest-variance specifically in the central millimeter. Nonetheless we can hypothesize that the thickest PR quartile encompasses mostly areas of the central millimeter, due to the high density of cone-type photoreceptor outer segments.^{45,46} Another hypothesis might suggest that mesopic testing conditions (as used in MAIA devices), mostly targeting twilight vision in rod dominated parafoveal areas. Therefore, the cone dominated central millimeter might not be stimulated in an optimal manner on MAIA examinations which may lead to the observed worse test-retest variance outcomes in the thickest PR quartile (I.Q).⁴⁷

Our study addresses pivotal findings on the relation between pointwise test-retest sensitivity outcomes and subclinical changes in PR-thickness. Nonetheless, due to limited patient size in our subgroups our findings should be confirmed further in larger patient cohorts, ideally in a multi-center setting. This also holds true when examining the effect of GA lesion size. We would hypothesize that with increase of GA lesion size the area of RPE and PR loss and subsequently test-retest variance increases. Based on overall GA area size (defined as RPE loss), the results were inconclusive (see supplemental document) which we primarily associate with our study design aimed at the comparison of pointwise changes.

The role of HRF is ambiguous as its presence related only to a slight increase of retest variance, which was only significant on examinations performed with the MP3. Due to their small size and high appearance rate in the junctional zone²⁴ their effect on retest variance

might be confounded with other previously discussed neurodegenerative changes characteristic of the area surrounding atrophic lesions. While the presence of drusen was associated with a slightly higher CoR compared to the areas without drusen, the difference was too small to be considered statistically significant. Still the detected higher variance can be explained by a reported decrease of photoreceptor sensibility over drusen.²⁹ Drusen thickness showed no clear trend regarding an increase or decrease of retest variances in our analysis. As a correlation between drusen regression and the emergence of atrophic lesions has been demonstrated¹⁹ an analysis of patients with an intermediate AMD might be better suited to determine the role of drusen alone on repeatability in MP measurements.

While both devices present a similar overall test-retest variance, the main aim of this study is the assessment of pointwise intradevice test-retest outcomes. Therefore, the results should not be used for a statement about interchangeability of the devices. Observed differences must be attributed to differing background luminosity as well as manufacturer specific system differences for stimulus projection. Previous studies have explored establishing a conversion formula between different microperimetry devices.^{48,49} Longitudinal studies including a larger patient cohort will be needed to establish a sound conversion methodology. Both devices presented an excellent average fixation stability according to the Fuji classification.⁵⁰ As both devices used an eye-tracking function foveal sparing had no significant impact on fixation stability.

The use of objective and reliable deep learning algorithms to quantify disease-specific imaging biomarkers stands as one of the principal strengths of this analysis. Given the continuously expanding amount of available information, the application of AI algorithms is set to become indispensable in the assessment of extensive imaging datasets. The prospective nature of the study, randomized device order, and standardized testing protocols further enhance the robustness of this analysis. The relatively small number of included patients must be considered a limitation of this study. Repeated measurements using the inbuilt follow-up as well as the registration of the stimuli points on the OCT might still encompass an error margin when analyzing the pointwise intra-device repeatability. Furthermore, due to repeated testing on the same day, a potential learning effect, as well as an effect of fatigue, regarding the length of the examinations, must be acknowledged as potential limitation. Meanwhile our analysis of average reaction time and mean retinal sensitivity between baseline and follow-up as well as repeatability between the first and second used device attributed no significance to

both effects. A potential bias might have arisen as patients were informed beforehand on the length and modalities of the study. Therefore, patients with a reduced compliance might have decided not to participate in this study.

The findings of this study strengthen the role of MP testing as a valuable functional endpoint in clinical trials and enable physicians in an everyday clinical setting to correctly interpret potential sensitivity loss observed in repeated MP measurements during long-term follow-up of GA patients. With the introduction of AI-based OCT analysis, providing for the first time a precise assessment of PR integrity loss and measures for distinct PR thinning, a functional topographic assessment becomes available. A higher test-retest variance within the area of PR integrity loss which exceeds the RPE loss border, must be accounted for during the analysis of disease progression, as PR thickness varies individually and over time. While mesopic MAIA and photopic MP3 testing have a similar retest variance, mesopic conditions might prove to provide more repeatable results considering the high variance within the atrophic lesion under photopic conditions.

Finally, with regards to future developments, this study highlights the merits of enhancing retinal sensitivity assessment with personalized, lesion-adapted grids, ideally created in a real-time manner. A special emphasis of future clinical trials examining functional impact of GA progression should be laid on the junctional zone, that we interpret as the highly variable area of perilesional PR loss, as this region will primarily be affected by disease progression.

Acknowledgements

A) Funding:

The study has not received any funding.

B) Competing interests and financial disclosures:

US-E is a scientific consultant for Apellis Pharmaceuticals, Bayer, EcoR1, AbbVie, Medscape, Johnson&Johnson, Allergan, Roche, Böhringer, Heidelberg, Novartis, Galimedix, Aviceda Therapeutics, Annexon Bioscience.

US-E received grant support from Genentech, Heidelberg Engineering, Kodiak, Novartis, Roche, RetInSight, Apellis Pharmaceuticals that were not related to this project.

GSR is a scientific consultant for Apellis, Bayer, Boehringer Ingelheim and Roche and received grant support from RetInSight that were not related to this project.

HB received research funds from Apellis, Heidelberg Engineering, RetInSight, Bayer and Roche that were not related to this project.

The other authors have nothing to declare.

Conflicts of interest: none

C): Other Acknowledgments

No other Acknowledgments

Authors' contributions:

LMC, GSR, KK are the principal investigator for the study. USE, GSR, PF, LMC, KK delineated the concept and design of the study. US-E and GSR were involved in the coordination of the study. AZ, LMC and PF were responsible for the acquisition of the data. SSW and MG were responsible for the Co-registration of microperimetry testings and OCT imaging. HB was responsible for the automated biomarker quantification. LMC and GSR performed the statistical analysis. LMC, KK, GSR, US-E, HB, MG, SSW and PF contributed to the generation of the manuscript. All authors reviewed and approved the final manuscript.

Ethics approval and consent to participate:

Ethics approval was obtained from the ethical commission of the Medical University Vienna in 2021(EK1399.2021)

Data availability statement

The data that support the findings of this study are available from the corresponding author SE, upon reasonable request.

References:

1. Bird AC, Phillips RL, Hageman GS. Geographic Atrophy: A Histopathological Assessment. *JAMA Ophthalmol.* 2014;132(3):338-345. doi:10.1001/JAMAOPHTHALMOL.2013.5799
2. Sadda SR, Guymer R, Holz FG, et al. Consensus Definition for Atrophy Associated with Age-Related Macular Degeneration on OCT: Classification of Atrophy Report 3. *Ophthalmology.* 2018;125(4):537-548. doi:10.1016/J.OPHTHA.2017.09.028
3. Boyer DS, Schmidt-Erfurth U, Van Lookeren Campagne M, Henry EC, Brittain C. THE PATHOPHYSIOLOGY OF GEOGRAPHIC ATROPHY SECONDARY TO AGE-RELATED MACULAR DEGENERATION AND THE COMPLEMENT PATHWAY AS A THERAPEUTIC TARGET. *Retina.* 2017;37(5):819. doi:10.1097/IAE.0000000000001392
4. Wong WL, Su X, Li X, et al. Global prevalence of age-related macular degeneration and disease burden projection for 2020 and 2040: a systematic review and meta-analysis. *Lancet Glob Health.* 2014;2(2):e106-16. doi:10.1016/S2214-109X(13)70145-1
5. Khan H, Aziz AA, Sulahria H, et al. Emerging Treatment Options for Geographic Atrophy (GA) Secondary to Age-Related Macular Degeneration. *Clin Ophthalmol.* 2023;17:321. doi:10.2147/OPTH.S367089
6. Sarda SP, Heyes A, Bektas M, et al. Humanistic and Economic Burden of Geographic Atrophy: A Systematic Literature Review. *Clin Ophthalmol.* 2021;15:4629-4644. doi:10.2147/OPTH.S338253
7. Apellis Announces FDA Acceptance of NDA Amendment and New PDUFA Date of February 26, 2023 for Pegcetacoplan for Geographic Atrophy (GA) - Apellis Pharmaceuticals, Inc. Accessed January 3, 2023. <https://investors.apellis.com/news-releases/news-release-details/apellis-announces-fda-acceptance-nda-amendment-and-new-pdufa>
8. Sadda SR, Chakravarthy U, Birch DG, Staurenghi G, Henry EC, Brittain C. Clinical Endpoints for the Study of Geographic Atrophy Secondary to Age-Related Macular Degeneration. *Retina.* 2016;36(10):1806. doi:10.1097/IAE.0000000000001283
9. Csaky KG, Patel PJ, Sepah YJ, et al. Microperimetry for geographic atrophy secondary to age-related macular degeneration. *Surv Ophthalmol.* 2019;64(3):353-364. doi:10.1016/j.survophthal.2019.01.014
10. Csaky K, Ferris F, Chew EY, Nair P, Cheetham JK, Duncan JL. Report From the NEI/FDA Endpoints Workshop on Age-Related Macular Degeneration and Inherited Retinal Diseases. *Invest Ophthalmol Vis Sci.* 2017;58(9):3456-3463. doi:10.1167/IOVS.17-22339
11. 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. doi:10.1016/J.PRETEYERES.2020.100907
12. Midena E, Pilotto E. Microperimetry in age: related macular degeneration. *Eye 2017* 31:7. 2017;31(7):985-994. doi:10.1038/eye.2017.34
13. Fleckenstein M, Mitchell P, Freund KB, et al. The Progression of Geographic Atrophy Secondary to Age-Related Macular Degeneration. *Ophthalmology.* 2018;125(3):369-390. doi:10.1016/J.OPHTHA.2017.08.038
14. Shen LL, Sun M, Khetpal S, Grossetta Nardini HK, Del Priore L V. Topographic Variation of the Growth Rate of Geographic Atrophy in Nonexudative Age-Related Macular Degeneration: A Systematic Review and Meta-analysis. *Invest Ophthalmol Vis Sci.* 2020;61(1). doi:10.1167/IOVS.61.1.2

15. Pfau M, Von Der Emde L, De Sisternes L, et al. Progression of Photoreceptor Degeneration in Geographic Atrophy Secondary to Age-related Macular Degeneration. *JAMA Ophthalmol*. 2020;138(10):1026-1034. doi:10.1001/JAMAOPHTHALMOL.2020.2914
16. Mai J, Riedl S, Reiter GS, et al. Comparison of FAF versus OCT-based evaluation of the therapeutic response to pegcetacoplan in Geographic Atrophy. *Am J Ophthalmol*. Published online July 16, 2022. doi:10.1016/J.AJO.2022.06.023
17. Riedl S, Vogl WD, Mai J, et al. The Effect of Pegcetacoplan Treatment on Photoreceptor Maintenance in Geographic Atrophy Monitored by Artificial Intelligence–Based OCT Analysis. *Ophthalmol Retina*. 2022;0(0). doi:10.1016/j.oret.2022.05.030
18. Schmetterer L, Scholl H, Garhöfer G, et al. Endpoints for clinical trials in ophthalmology. *Prog Retin Eye Res*. Published online January 2, 2023:101160. doi:10.1016/J.PRETEYERES.2022.101160
19. Schlanitz FG, Baumann B, Kundi M, et al. Drusen volume development over time and its relevance to the course of age-related macular degeneration. *Br J Ophthalmol*. 2017;101(2):198. doi:10.1136/BJOPHTHALMOL-2016-308422
20. Bui PTA, Reiter GS, Fabianska M, et al. Fundus autofluorescence and optical coherence tomography biomarkers associated with the progression of geographic atrophy secondary to age-related macular degeneration. *Eye* 2021. Published online August 16, 2021:1-7. doi:10.1038/s41433-021-01747-z
21. Biarnés M, Garrell-Salat X, Gómez-Benlloch A, et al. Methodological Appraisal of Phase 3 Clinical Trials in Geographic Atrophy. *Biomedicines*. 2023;11(6). doi:10.3390/BIOMEDICINES11061548
22. Lachinov D, Seeböck P, Mai J, Goldbach F, Schmidt-Erfurth U, Bogunovic H. Projective Skip-Connections for Segmentation Along a Subset of Dimensions in Retinal OCT. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. 2021;12901 LNCS:431-441. doi:10.1007/978-3-030-87193-2_41
23. Vogl WD, Riedl S, Mai J, et al. Predicting Topographic Disease Progression and Treatment Response of Pegcetacoplan in Geographic Atrophy Quantified by Deep Learning. *Ophthalmol Retina*. 2022;0(0). doi:10.1016/j.oret.2022.08.003
24. Schmidt-Erfurth U, Bogunovic H, Grechenig C, et al. Role of Deep Learning-Quantified Hyperreflective Foci for the Prediction of Geographic Atrophy Progression. *Am J Ophthalmol*. 2020;216:257-270. doi:10.1016/J.AJO.2020.03.042
25. Mai J, Lachinov D, Riedl S, et al. Clinical validation for automated geographic atrophy monitoring on OCT under complement inhibitory treatment. *Scientific Reports* 2023 13:1. 2023;13(1):1-11. doi:10.1038/s41598-023-34139-2
26. Orlando JI, Gerendas BS, Riedl S, et al. Automated Quantification of Photoreceptor alteration in macular disease using Optical Coherence Tomography and Deep Learning. *Scientific Reports* 2020 10:1. 2020;10(1):1-12. doi:10.1038/s41598-020-62329-9
27. Asgari R, Orlando JI, Waldstein S, et al. Multiclass segmentation as multitask learning for drusen segmentation in retinal optical coherence tomography. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. 2019;11764 LNCS:192-200. doi:10.1007/978-3-030-32239-7_22
28. Schlegl T, Bogunovic H, Klimescha S, et al. Fully Automated Segmentation of Hyperreflective Foci in Optical Coherence Tomography Images. Published online May 8, 2018. Accessed October 12, 2023. <https://arxiv.org/abs/1805.03278v1>
29. Reiter GS, Bogunovic H, Schlanitz F, et al. Point-to-point associations of drusen and hyperreflective foci volumes with retinal sensitivity in non-exudative age-related

- macular degeneration. *Eye* 2023. Published online May 11, 2023:1-7.
doi:10.1038/s41433-023-02554-4
30. Arikan M, Sadeghipour A, Gerendas B, Told R, Schmidt-Erfurt U. Deep learning based multi-modal registration for retinal imaging. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. 2019;11797 LNCS:75-82. doi:10.1007/978-3-030-33850-3_9
 31. Told R, Reiter GS, Orsolya A, et al. SWEPT SOURCE OPTICAL COHERENCE TOMOGRAPHY ANGIOGRAPHY, FLUORESCEIN ANGIOGRAPHY, AND INDOCYANINE GREEN ANGIOGRAPHY COMPARISONS REVISITED: Using a Novel Deep-Learning-Assisted Approach for Image Registration. *Retina*. 2020;40(10):2010-2017. doi:10.1097/IAE.0000000000002695
 32. Martin Bland J, Altman DG. STATISTICAL METHODS FOR ASSESSING AGREEMENT BETWEEN TWO METHODS OF CLINICAL MEASUREMENT. *The Lancet*. 1986;327(8476):307-310. doi:10.1016/S0140-6736(86)90837-8
 33. Koo TK, Li MY. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *J Chiropr Med*. 2016;15(2):155. doi:10.1016/J.JCM.2016.02.012
 34. Pfau M, Lindner M, Müller PL, et al. Effective Dynamic Range and Retest Reliability of Dark-Adapted Two-Color Fundus-Controlled Perimetry in Patients With Macular Diseases. *Invest Ophthalmol Vis Sci*. 2017;58(6):BIO158-BIO167. doi:10.1167/IOVS.17-21454
 35. Heier JS, Lad EM, Holz FG, et al. Pegcetacoplan for the treatment of geographic atrophy secondary to age-related macular degeneration (OAKS and DERBY): two multicentre, randomised, double-masked, sham-controlled, phase 3 trials. *The Lancet*. 2023;402(10411):1434-1448. doi:10.1016/S0140-6736(23)01520-9
 36. Holz FG, Sadda SR, Busbee B, et al. Efficacy and Safety of Lampalizumab for Geographic Atrophy Due to Age-Related Macular Degeneration: Chroma and Spectri Phase 3 Randomized Clinical Trials. *JAMA Ophthalmol*. 2018;136(6):666-677. doi:10.1001/JAMAOPHTHALMOL.2018.1544
 37. Coulibaly LM, Mohamed H, Fuchs P, Schmidt-Erfurth U, Reiter GS. Inter and intradevice assessment of microperimetry testing in aging eyes. *Scientific Reports* 2024 14:1. 2024;14(1):1-9. doi:10.1038/s41598-024-51539-0
 38. Alibhai AY, Mehta N, Hickson-Curran S, et al. Test-retest variability of microperimetry in geographic atrophy. *Int J Retina Vitreous*. 2020;6(1):1-6. doi:10.1186/S40942-020-00217-0/FIGURES/2
 39. Pfau M, Müller PL, Von Der Emde L, et al. MESOPIC and DARK-ADAPTED TWO-COLOR FUNDUS-CONTROLLED PERIMETRY in GEOGRAPHIC ATROPHY SECONDARY to AGE-RELATED MACULAR DEGENERATION. *Retina*. 2020;40(1):169-180. doi:10.1097/IAE.0000000000002337
 40. Pfau M, Von Der Emde L, Dysli C, et al. Light Sensitivity Within Areas of Geographic Atrophy Secondary to Age-Related Macular Degeneration. *Invest Ophthalmol Vis Sci*. 2019;60(12):3992-4001. doi:10.1167/IOVS.19-27178
 41. Hariri AH, Tepelus TC, Akil H, Nittala MG, Sadda SR. Retinal Sensitivity at the Junctional Zone of Eyes With Geographic Atrophy Due to Age-Related Macular Degeneration. *Am J Ophthalmol*. 2016;168:122-128. doi:10.1016/J.AJO.2016.05.007
 42. Table View | A Study to Compare the Efficacy and Safety of Intravitreal APL-2 Therapy With Sham Injections in Patients With Geographic Atrophy (GA) Secondary to Age-Related Macular Degeneration | ClinicalTrials.gov. Accessed October 29, 2023. <https://clinicaltrials.gov/study/NCT03525613?tab=table>

43. Bird AC, Phillips RL, Hageman GS. Geographic Atrophy: A Histopathological Assessment. *JAMA Ophthalmol.* 2014;132(3):338. doi:10.1001/JAMAOPHTHALMOL.2013.5799
44. Zanzottera EC, Ach T, Huisingh C, Messinger JD, Spaide RF, Curcio CA. Visualizing retinal pigment epithelium phenotypes in the transition to geographic atrophy in age-related macular degeneration. *Retina.* 2016;36(Suppl 1):S12. doi:10.1097/IAE.0000000000001276
45. Curcio CA, Sloan KR, Kalina RE, Hendrickson AE. Human photoreceptor topography. *Journal of Comparative Neurology.* 1990;292(4):497-523. doi:10.1002/CNE.902920402
46. Wang YX, Pan Z, Xue CC, Xie H, Wu X, Jonas JB. Macular outer nuclear layer, ellipsoid zone and outer photoreceptor segment band thickness, axial length and other determinants. *Scientific Reports 2023 13:1.* 2023;13(1):1-10. doi:10.1038/s41598-023-32629-x
47. Simunovic MP, Moore AT, MacLaren RE. Selective Automated Perimetry Under Photopic, Mesopic, and Scotopic Conditions: Detection Mechanisms and Testing Strategies. *Transl Vis Sci Technol.* 2016;5(3):10-10. doi:10.1167/TVST.5.3.10
48. Liu H, Bittencourt MG, Wang J, et al. Assessment of Central Retinal Sensitivity Employing Two Types of Microperimetry Devices. *Transl Vis Sci Technol.* 2014;3(5):3-3. doi:10.1167/TVST.3.5.3
49. Balasubramanian S, Uji A, Lei J, Velaga S, Nittala M, Sadda SV. Interdevice comparison of retinal sensitivity assessments in a healthy population: the CenterVue MAIA and the Nidek MP-3 microperimeters. *British Journal of Ophthalmology.* 2018;102(1):109-113. doi:10.1136/BJOPHTHALMOL-2017-310258
50. Fujii GY, De Juan E, Sunness J, Humayun MS, Pieramici DJ, Chang TS. Patient selection for macular translocation surgery using the scanning laser ophthalmoscope. *Ophthalmology.* 2002;109(9):1737-1744. doi:10.1016/S0161-6420(02)01120-X

Figure Captions:

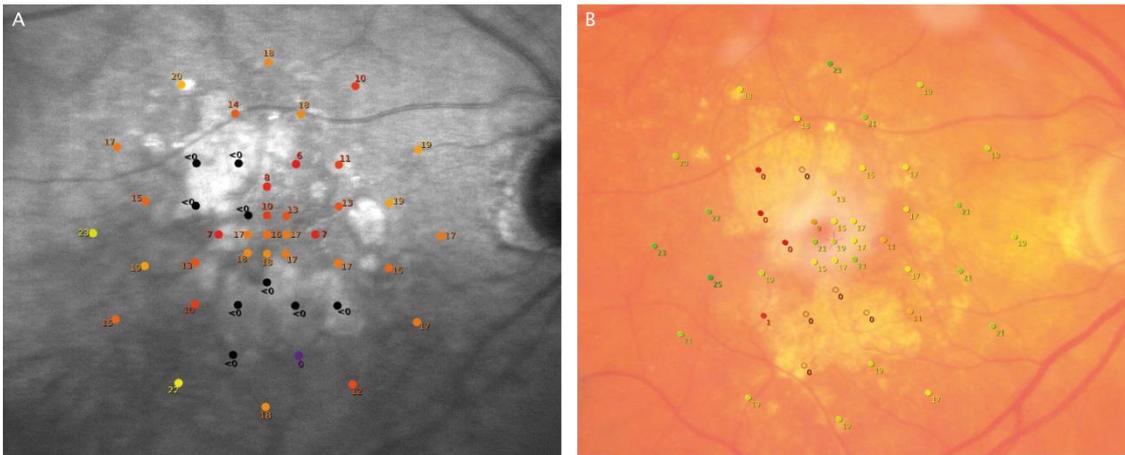


Figure 1: 45 stimuli study grid for MP examinations of a patient with geographic atrophy using CenterVue (iCare) MAIA (A) and Nidek MP3 (B)

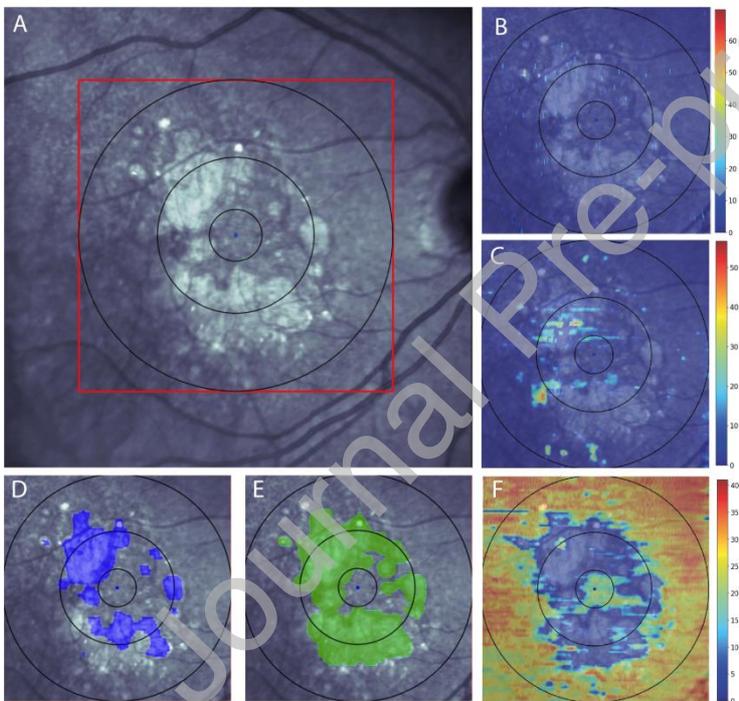


Figure 2: Automated imaging biomarker segmentation of a patient with a foveal sparing geographic atrophy. A) Near-infrared en-face image; B) Hyperreflective foci thickness map (scale in μm); C) Drusen thickness map (scale in μm); D) En-face 2D-representation of retinal pigment epithelium loss in blue; E) En-Face 2D-representation of photoreceptor integrity loss in green; F) Photoreceptor thickness map (scale in μm)

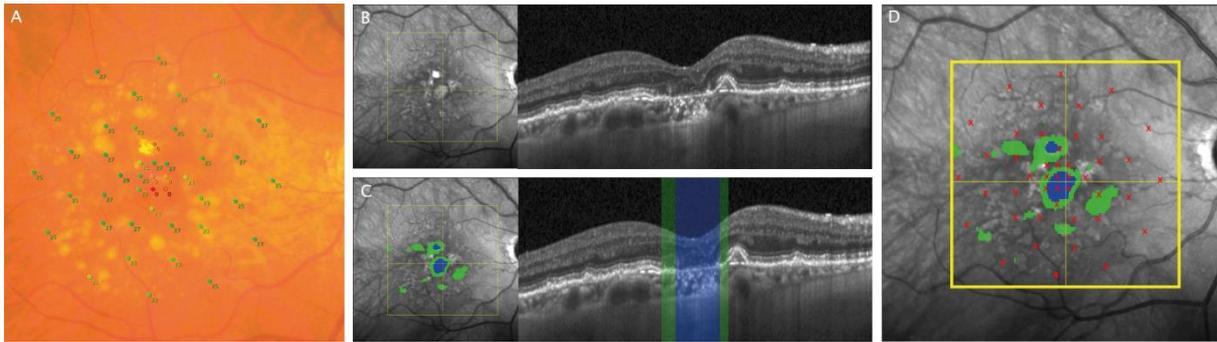


Figure 3: Figure 3: MP3 results with 45-stimuli grid on patient without foveal sparing geographic atrophy (A); Heidelberg Spectralis Near-infrared (NIR)-en-face image and Central B-Scan without (B) and with retinal pigment epithelium (RPE, blue) and photoreceptor integrity loss (PR, green) segmentation (C); Co-registration of microperimetry stimuli (marked by a red cross) on the NIR en-face image with 2 Dimensional RPE (blue) and PR loss (green) segmentation (D)

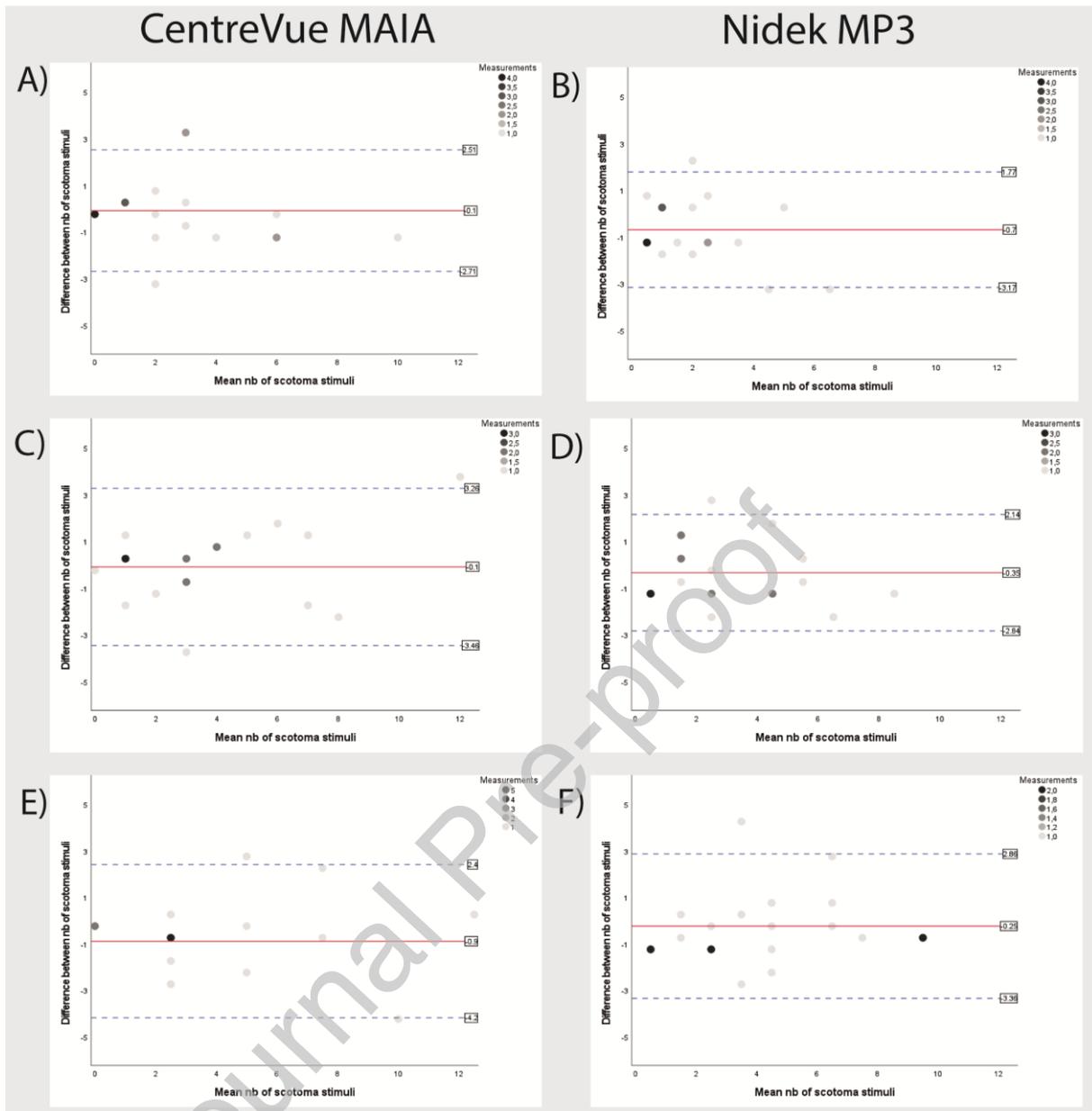


Figure 4: Figure 4. Bland-Altman plots for repeatability of scotoma detection in MAIA (left) and MP3 (right) defining scotoma as stimuli with PWS values ≤ 0 db (A and B), ≤ 5 db (C and D) and ≤ 10 db (E and F). Limits of Agreements are marked by the dotted line in blue.

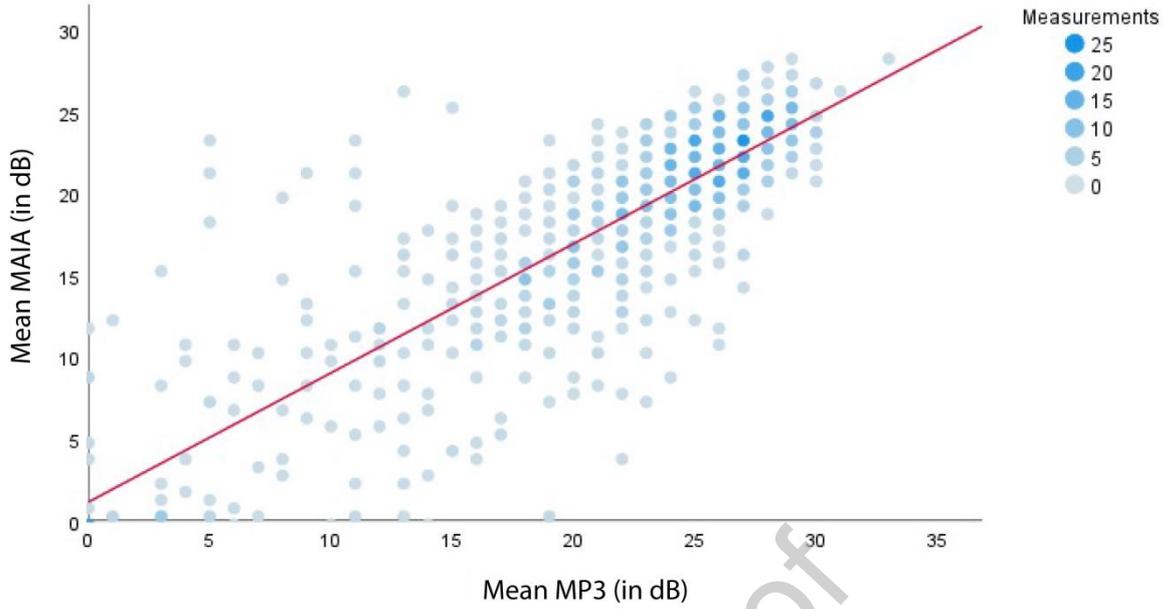


Figure 5: Figure 5. Graphical representation of linear regression model between mean pointwise sensitivity measurements on MAIA and MP3 with the linear regression curve in red. Measurements are grouped and color-coded as seen in the table on the left.

Table 1. Repeatability of Microperimetry

Group	Nb of Stimuli	MAIA (dB)					MP-3 (dB)					
		Mean PWS (SD)	SD	CoR	Upper LoA	Lower LoA	Nb of Stimuli	Mean PWS (SD)	SD	CoR	Upper LoA	Lower LoA
Per device overall	900	18.2 (6.7)	3.35	± 6.57	6.33	-6.88	900	21.6 (7.1)	3.36	± 6.59	6.2	-6.98
Foveal sparing	630 (70%)	18.1 (6.5)	3.25	± 6.38	5.78	-6.98	630 (70%)	18.1 (6.5)	3.41	± 6.68	6.06	-7.31
No foveal sparing	270 (30%)	18.4 (7.1)	3.47	± 6.81	7.18	-6.42	270 (30%)	18.4 (7.1)	3.17	± 6.21	6.4	-6.04
No RPE/PR loss	611 (67.9%)	20.6 (4.4)	3.07	$\pm 6.01^*$	5.36	-6.66	622 (69.1%)	24.7 (3.4)	2.61	$\pm 5.12^*$	4.5	-5.74
Only PR loss	136 (15.1%)	15.2 (6.3)	3.87	$\pm 7.59^*$	7.90	-7.27	139 (15.4%)	19.3 (5.6)	3.18	$\pm 6.24^*$	5.68	-6.8
RPE + PR loss	138 (15.3%)	10.2 (8.2)	3.82	$\pm 7.48^*$	8.04	-6.92	138 (15.3%)	9.9 (8.1)	5.6	$\pm 10.99^*$	11.81	-10.14
RPE loss	138 (15.6%)	10.2 (8.2)	3.82	± 7.48	8.04	-6.92	138 (15.5%)	9.9 (8.1)	5.6	$\pm 10.99^*$	11.81	-10.15
No RPE loss	747 (84.4%)	19.6 (5.2)	3.25	± 6.37	5.89	-6.84	761 (84.5%)	23.7 (4.4)	2.72	$\pm 5.34^*$	4.73	-5.95
PR loss	274 (30.9%)	12.7 (7.7)	3.81	$\pm 7.52^*$	7.96	-7.08	277 (30.8%)	14.6 (8.4)	4.6	$\pm 9.01^*$	9.14	-8.88
No PR loss	611 (69.1%)	20.6 (4.4)	3.07	$\pm 6.01^*$	5.36	-6.66	622 (69.2%)	24.7 (3.4)	2.61	$\pm 5.12^*$	4.5	-5.74
Presence of HRF	73 (8.3%)	14.8 (8.1)	3.31	± 6.49	6.49	-6.49	86 (9.6%)	17 (8.9)	4.7	$\pm 9.21^*$	9.5	-8.92
Absence of HRF	812 (91.7%)	18.4 (6.5)	3.37	± 6.6	5.89	-6.85	813 (90.4%)	22.08 (6.8)	3.19	$\pm 6.25^*$	5.79	-6.71
Presence of drusen	512 (57.8%)	17.4 (7.1)	3.45	± 6.77	6.52	-7.02	546 (60.7%)	21.3 (7)	3.39	± 6.64	6.08	-7.2
Absence of drusen	373 (42.1%)	19.2 (6)	3.27	± 6.33	5.93	-6.73	353 (39.3%)	22.1 (7.3)	3.33	± 6.53	5.97	-7.09

Table 1 Test-retest repeatability according to geographic atrophy imaging biomarker; Subgroups with heterogenous variances are marked by an *

Table 2. Repeatability according to Photoreceptor and Drusen thickness

Quartiles	MAIA (dB)						MP-3 (dB)					
	Nb of stimuli	Thickness range	SD	CoR	Upper LoA	Lower LoA	Nb of stimuli	Thickness range	SD	CoR	Upper LoA	Lower LoA
PR thickness I. Q	225	>29.12	3.78	± 7.41	7.46	-7.36	225	>29.38	2.72	$\pm 5.33^*$	4.64	-6.02
PR thickness II. Q	224	29.12 – 24.53	2.98	$\pm 5.84^*$	5.39	-6.29	224	29.38 – 24.59	2.71	$\pm 5.31^*$	4.93	-5.69
PR thickness III. Q	225	24.53 – 10.97	2.86	$\pm 5.6^*$	5.14	-6.06	225	24.59 – 11.52	2.68	$\pm 5.26^*$	4.61	-5.91
PR thickness IV. Q	225	<10.97	3.68	$\pm 7.2^*$	6.85	-7.55	225	<11.52	4.8	$\pm 9.41^*$	8.95	-9.86
Drusen thickness I. Q	127	>16.68	3.14	± 6.16	6.13	-6.19	136	>15.49	3.04	± 5.96	5.41	-6.51
Drusen thickness II. Q	129	16.68 – 7.71	3.42	± 6.71	6.18	-7.24	137	15.49 – 6.31	5.51	± 6.89	6.06	-7.71
Drusen thickness III. Q	128	7.71 – 2.92	4	± 7.85	7.74	-7.95	137	6.31 – 2.65	3.1	± 6.08	6.25	-5.91
Drusen thickness IV. Q	127	<2.92	3.09	± 6.06	5.78	-6.34	136	<2.65	3.76	± 7.38	7.47	-7.28

Table 2. Test-retest repeatability according to photoreceptor and drusen thickness quartiles, Subgroups with heterogeneous variances are marked by an *.

Table 3: Interclass correlation coefficients

	MAIA I	MAIA II	MP-3 I	MP-3 II
MAIA I	1	0.94 [0.93-0.95]	0.81 [0.55-0.9]	0.83 [0.45-0.92]
MAIA II	0.94 [0.93-0.95]	1	0.84 [0.62-0.91]	0.86 [0.5-0.94]
MP-3 I	0.81 [0.55-0.9]	0.84 [0.62-0.91]	1	0.94 [0.94-0.95]
MP-3 II	0.83 [0.45-0.92]	0.86 [0.5-0.94]	0.94 [0.94-0.95]	1

Table 3. Interclass correlation coefficients between each performed microperimetry examination

Table 4. Fixation Stability

	2° (P1)		4° (P2)	
	Median (IQR)	Effect size	Median (IQR)	Effect size
MAIA	87.3 (12.1)	2.62 (p=0.48)	98.05 (4.3)	1.42 (p=0.49)
MP3	90 (10)		98.5 (4.8)	
Run 1	87.3 (11.3)	0.42 (p=0.98)	97 (5.1)	-1.1 (p=0.59)
Run 2	90.2 (11.3)		98.2 (3.3)	
Foveal sparing	87.67 (12.1)	3.99 (p=0.33)	98.1 (4.2)	3.29 (p=0.14)
No foveal sparing	89 (9.9)		98.05 (4.4)	

Table 4. Results for fixation stability in central 2° and 4° according to device, first or second examination (Run), presence or absence of foveal sparing

Table of Contents Statement

This analysis provides much needed test-retest repeatability references regarding pointwise retinal sensitivity assessed with microperimetry in patients with geographic atrophy. Artificial intelligence-based quantifications of disease specific optical coherence tomography biomarkers enabled a comprehensive analysis based on underlying morphologic retinal alterations. Retinal pigment epithelium and/or photoreceptor loss was associated with higher retest variances. These findings are crucial regarding everyday patient monitoring as well as the design of future clinical trials evaluating disease progression.

Competing interests and financial disclosures:

US-E is a scientific consultant for Apellis Pharmaceuticals, Bayer, EcoR1, AbbVie, Medscape, Johnson&Johnson, Allergan, Roche, Böhringer, Heidelberg, Novartis, Galimedix, Aviceda Therapeutics, Annexon Bioscience.

US-E received grant support from Genentech, Heidelberg Engineering, Kodiak, Novartis, Roche, RetInSight, Apellis Pharmaceuticals that were not related to this project.

GSR is a scientific consultant for Bayer and received grant support from RetInSight that were not related to this project.

HB received research funds from Apellis, Heidelberg Engineering, RetInSight, Bayer and Roche that were not related to this project.

The other authors have nothing to declare.

Conflicts of interest: none