AMERICAN ACADEMY OF OPHTHALMOLOGY®

Customized Evaluation of Progressive Visual Sensitivity Loss in Geographic Atrophy to Improve the Power of Clinical Trials

Abera Saeed, MChD,[1,2] *Robyn H. Guymer, MBBS, PhD,*[1,2] *Xavier Hadoux, MEng, PhD,*[1,2] *Maxime Jannaud, MEng,*[1] *Darvy Dang, BOrth(Hons),*[1] *Lauren A.B. Hodgson, MPH,*[1] *Emily K. Glover, OD,*[1] *Erin E. Gee, BAppSc(MedRad),*[1] *Peter van Wijngaarden, MBBS(Hons), PhD,*[1,2] *Zhichao Wu, BAppSc(Optom), PhD*[1,2]

Purpose: To evaluate the effectiveness of different approaches for customizing the selection of a subset of test locations on defect-mapping microperimetry (DMP) for improving the detection of progressive visual function decline in geographic atrophy (GA).

Design: Prospective longitudinal study.

Participants: Sixty eyes from 53 participants with GA secondary to age-related macular degeneration.

Methods: Participants underwent 3-monthly DMP testing twice at each visit for up to 24 months, where the extent of deep visual sensitivity losses on each test was determined through single presentations of 10-decibel stimuli at 208 locations within the central 8° radius region. Seven outcome measures were derived, which included evaluating the proportion of locations missed (PLM; showing nonresponse to stimuli) on DMP in a subset of test locations based on their proximity to the GA margin, or to locations neighboring repeatably nonresponding points on 2 baseline tests (i.e., missed both tests at baseline). These outcome measures were compared by their coefficient of variation (CoV; reflecting performance for capturing longitudinal changes) and sample size estimates in a 2-arm trial seeking to detect a ≥30% treatment effect. Changes in GA extent and best-corrected visual acuity (BCVA) were evaluated for comparison.

Main Outcome Measures: Coefficient of variation and sample size estimates.

Results: Evaluating PLM at points immediately adjacent (<1°) to repeatably nonresponding test locations at baseline (CoV = 47%) was the best performing outcome measure on DMP testing. This measure outperformed BCVA (CoV = 188%; $P < 0.001$) at detecting longitudinal changes and was comparable to evaluating GA extent (CoV = 58%; $P = 0.097$). Sample size requirements in a 24-month trial using this outcome measure on DMP testing were lower by 46% and 94% compared with evaluating GA extent and BCVA, respectively.

Conclusions: Customized evaluation of DMP functional testing results in regions adjacent to repeatably nonresponding locations at baseline improved the detection of longitudinal changes compared with the evaluation of all test locations. These findings show that it is possible to sensitively capture progressive visual function decline with this approach, supporting its use in future GA treatment trials.

Financial Disclosure(s): Proprietary or commercial disclosure may be found in the Footnotes and Disclosures at the end of this article. *Ophthalmology Science 2025;5:100763*

Geographic atrophy (GA) is a late stage of age-related macular degeneration (AMD) characterized by the progressive loss of the retinal pigment epithelium that is accompanied by degeneration of the overlying photoreceptors.[1–5] Eyes with GA exhibit progressive visual function decline,[6–10] and therefore a key goal of GA treatment is to ultimately slow or prevent the worsening of such visual function loss.

Recent landmark trials have shown that treatments with intravitreal complement inhibitors have a significant effect on slowing GA growth.[11,12] However, a significant treatment effect was not observed during the trial period in any of the prespecified visual function outcome measures, including best-corrected visual acuity (BCVA), reading speed, and visual sensitivity on microperimetry (or fundus-tracked perimetry).[11,12] The European Medicines Agency has previously expressed a desire to see evidence of treatment benefits on GA growth supported by at least a trend of a positive effect on visual function outcome measures when considering the approval of such treatments.[13,14] Although the US Food and Drug Administration has expressed an acceptance of GA growth alone as an endpoint,[13,15] the US Food and Drug Administration has expressed a clear preference for visual function over anatomic outcome measures.[15] Thus, establishing sensitive methods for capturing progressive visual function decline in eyes with GA is invaluable for the evaluation of new interventions.

https://doi.org/10.1016/j.xops.2025.100763
ISSN 2666-9145/25

We have recently developed an approach optimized to capture the spatial extent of deep visual sensitivity losses in GA, termed "defect-mapping" microperimetry (DMP) testing.[16,17] Unlike standard threshold-based microperimetry, which presents multiple stimuli at a limited number of test locations, DMP testing allows for higher spatial density testing through single presentations of suprathreshold stimuli over a similar test duration.[16] We have recently shown that the global measure of the proportion of locations missed (PLM) on DMP testing showed a strong correlation with GA extent in the corresponding region tested at cross section (Spearman rank correlation coefficient = 0.90),[17] and that its longitudinal changes also showed moderate correlations with change in GA extent (Spearman rank correlation coefficient = 0.52).[18] These findings are notably different compared with previous findings where the widely used measure of mean sensitivity on standard threshold-based microperimetry showed only moderate cross-sectional and weak longitudinal correlations with GA extent (Spearman rank correlation coefficient = −0.42 at baseline and approximately 0.01 at 48 weeks from baseline, respectively).[10] This suggests that DMP testing could potentially serve as a more sensitive visual function outcome measure in GA trials compared with threshold-based microperimetry testing.

Furthermore, a recent study showed that evaluation of specific subsets of test locations in standard threshold-based microperimetry testing may offer greater sensitivity in detecting progressive visual sensitivity decline compared with the standard approach of evaluating all test locations.[19] Specifically, this study demonstrated that evaluating changes in mean visual sensitivity of test locations adjacent to absolute scotomas (or at "perilesional" locations), or all nonscotomatous locations, at baseline was more sensitive than evaluating changes in mean visual sensitivity at all test locations.[19] Similarly, post hoc analysis of a recent landmark GA trial showed treatment benefit on reducing the number of absolute scotomas for test locations at the junctional zone of atrophy (or points within ±250 μm of the GA margin).[11] Although this result reached statistical significance for only one treatment group when using standard threshold-based testing, these findings underscore the potential of customized evaluation of a subset of test locations on DMP for detecting beneficial treatment effects. Thus, this study sought to examine different approaches for customized evaluation of DMP testing results and their implications for improving the power of future GA trials for detecting beneficial treatment effects on a visual function outcome.

Methods

This study included individuals enrolled in a prospective study of atrophic AMD undertaken at the Centre for Eye Research Australia, which was conducted in accordance with the International Conference on Harmonization Guidelines for Good Clinical Practice and the tenets of the Declaration of Helsinki. Institutional Review Board approval was obtained for this study from the Royal Victorian Eye and Ear Hospital, and all participants provided informed consent. Participants in this study were enrolled between February 2019 and September 2023, an extended period due to the impact of the coronavirus 2019 pandemic.

Participants

This study enrolled individuals ≥50 years with GA secondary to AMD, which was defined as the presence of ≥1 well-demarcated area of decreased autofluorescence that was ≥175 μm in diameter on fundus autofluorescence (FAF) imaging.[20] Such GA had to be present in ≥1 eye with a BCVA of 20/100 or better and without neovascular AMD or a history of any intervention for AMD. Individuals with any ocular, systemic, or neurologic condition(s) apart from GA secondary to AMD that prohibits the reliable assessment of the retina or visual function were excluded from this study.

This study sought to perform microperimetry testing in these participants at 3-monthly intervals, up to a maximum of 9 visits in total. If both eyes were deemed eligible at baseline, participants either performed testing of only 1 eye or both eyes at all follow-up visits or both eyes at different follow-up visits (such as longitudinal testing of 1 eye over a 24-month period and then longitudinal testing of the other eye over another 24-month period). The schedule of follow-up testing was chosen based on the preference of the participant, and only visits within 24 months from baseline were included in the analyses of this study.

Retinal Imaging

Fundus autofluorescence images were obtained using the Spectralis HRA+OCT device (Heidelberg Engineering GmbH), using a 488-nm excitation light source and capturing emitted fluorescence signals between 500 and 700 nm. Fundus autofluorescence images of the central 30° × 30° region were obtained, having a minimum dimension of 768 × 768 pixels and with ≥100 frames averaged. OCT volume scans of the central 20° × 20° were also obtained using the same device and consisted of 97 B-scans, each with 1024 A-scans and 16 frames averaged. Color fundus photographs of the central 45° diameter were also obtained using the Canon CR6-45NM device (Canon).

Microperimetry Testing

Microperimetry testing was performed using the Macular Integrity Assessment device (CenterVue), after pupillary dilation and before any assessments that could bleach the retina or compromise the ocular surface. Details about the specifications of this device have been described in detail previously.[21,22] Briefly, Goldmann size III (0.43° diameter) stimuli with a luminance between 1.35 and 318 cd/m^2 can be presented against a mesopic background (1.27 cd/m^2) with this device, which provides a dynamic range of 36 decibels (dB) of differential contrast. In this study, the DMP testing strategy[16,17] was used to assess the spatial extent of deep visual sensitivity abnormalities. This involved single presentations of only stimuli with a luminance of 32.9 cd/m^2 (or 10 dB, relative to the dynamic range of this device), once at each of 208 locations using an isotopic stimulus pattern with an interstimulus interval of 1° in the central 8° radius region (central 4.6 mm diameter); these tests took approximately 6 minutes, as reported previously.[17] This stimulus luminance used for DMP testing corresponds to the floor of the effective dynamic range of this device (previously defined as the highest luminance where 5% of its retested values are ≤0 dB, effectively being statistically indistinguishable from the floor of the physical dynamic range of the device).[23] This stimulus grid was manually centered on the anatomic fovea if the fixation was nonfoveal, and a 1° diameter central red fixation target was used during testing.

Participants were required to complete 2 reliable DMP tests at each visit, which was defined as a test with ≤25% false-positive errors (determined through catch trials of 10-dB stimulus

presentations to the optic nerve head), and unreliable tests were discarded and repeated where possible. A few minutes of rest were always offered to the participants between each DMP test to minimize potential fatigue effects. Note that we previously reported that there was no evidence of a significant learning effect with DMP testing,[17] which is in contrast to what we have observed previously with standard threshold-based microperimetry testing.[21] After the first DMP test, all subsequent tests (including those during the same visit or at subsequent visits) were performed using the "follow-up" function so that the same locations in the retina are assessed across all tests.

Image Processing and Annotations

Fundus autofluorescence images from the follow-up visits were all manually coregistered with the baseline FAF image through radial and affine transformations[24] of the images after the placement of fiducial points at retinal blood vessel bifurcations on the FAF images. The near-infrared fundus image captured by the microperimeter at baseline was also manually coregistered with the baseline FAF image, using the near-infrared fundus image with the best quality from the 2 DMP tests at baseline. Because all microperimetry tests were performed using the "follow-up" function, these abovementioned steps result in the coregistration of the FAF and DMP tests at all visits in this study.

Geographic atrophy on the FAF images (i.e., regions that were $\geq$175 μm in diameter)[25] were annotated by a grader (D.D. or E.E.G.) and then reviewed together with a senior grader (L.A.B.H. or Z.W.) while being masked to the microperimetry testing results. Any disagreements were resolved through open adjudication and revision of the annotations together. These annotations were performed using custom-written software (Cross-Modality Annotation Software; Ophthalmic Neuroscience Unit, Centre for Eye Research Australia).[26,27]

Deriving Microperimetry Outcome Measures

Figure 1 illustrates 7 different DMP outcome measures that were derived for comparison in this study. This included evaluating the PLM or showing nonresponse to a presented stimuli on DMP in a subset of points selected based on their: (1) proximity to the GA margin on FAF imaging at baseline or (2) proximity to locations that were repeatably nonresponding on the 2 baseline DMP tests (i.e., were missed on both tests), as follows.

1. Measure 1: All locations—including all points tested.
2. Measure 2: Junctional zone—including points within ±250 μm of the GA margin.
3. Measure 3: Perilesional zone—including points between 0-250 μm outside the GA margin.
4. Measure 4: Perilesional zone—including points between 0-500 μm outside the GA margin.
5. Measure 5: Peri-scotomatous zone—including points $\leq$1° neighboring repeatably nonresponding locations at baseline (along the horizontal or vertical axis).
6. Measure 6: Peri-scotomatous zone—including points $\leq$2° neighboring repeatably nonresponding locations at baseline (along the horizontal or vertical axis).
7. Measure 7: Nonscotomatous zone—including all points excluding repeatably nonresponding locations.

Statistical Analyses

Longitudinal changes from baseline at each 3-monthly follow-up visit up to 24 months for the 7 measures of PLM on DMP above were evaluated using random intercepts models. These models included random intercepts at the individual, eye, and visit level to account for correlations between 2 eyes of an individual, between multiple visits from the same eye, and between the 2 DMP tests at each visit. Similarly, changes in the percentage of the central 8° region with GA and BCVA at each follow-up visit were also evaluated using a random intercepts model but only including random intercepts at the individual and eye level (as each of these parameters was measured only once per visit).

The performance of the different DMP outcome measures for capturing longitudinal changes was then evaluated and compared by calculating their coefficient of variation (CoV),[19] a normalized measure of between-individual variability relative to the change captured. The CoV was derived by dividing the standard deviation of the rate of PLM change over time (per year) by its mean, and these 2 parameters were derived from a linear mixed model including random intercepts and slopes at the individual and eye level and random intercepts at the visit level. The CoV for the percentage of the central 8° region with GA and BCVA was similarly derived for comparison with the DMP outcome measures because these are 2 of the most commonly used outcome measures in GA trials.[28] Note that the CoV is presented as a percentage, with lower values indicating better performance for detecting differences in longitudinal changes between groups (e.g., treatment arms in an interventional trial). The 95% confidence intervals of the CoV for each outcome measure, and the statistical significance of the difference between each outcome measure, were derived with a nonparametric bootstrap resampling procedure (n = 1000 resamples at the individual level). These analyses were all performed using Stata software version 18 (StataCorp).

Sample size requirements for clinical trials seeking to detect a difference in the different PLM outcome measures above, GA extent, and BCVA were calculated using the estimates derived from the linear mixed model described earlier. Sample size estimates were determined for a 2-arm clinical trial seeking to detect a $\geq$30% treatment effect with 80% power and a 2-sided α of 0.05, assuming a 10% attrition rate per year. These estimates were calculated separately for a follow-up period of 12 and 24 months assuming 6-monthly visits, assuming 2 visits at baseline, using the "*longpower*" package in R Statistical Software (R Core Team, 2021).[29]

Results

This study included 60 eyes from 53 participants who were on average 76 ± 7 years (range, 52−87 years) and 31 (62%) participants were female. The median total GA area of these eyes at baseline was 3.81 mm^2 (interquartile range [IQR], 2.39−9.37 mm^2) or 1.95 mm (IQR, 1.54−3.06 mm) for square-root GA area. The median percentage of GA in the central 8° radius region sampled by DMP testing was 20% (IQR, 13%−43%). These eyes were seen across a median of 5 visits (IQR, 4 to 7 visits), over a median follow-up of 15 months (IQR, 12−20 months; 85% of eyes had a follow-up duration of $\geq$12 months). The median rate of GA growth was 1.93 mm^2/year (IQR, 1.06−3.66 mm^2/year) or 0.40 mm/year (IQR, 0.27−0.63 mm/year) for square-root GA area.

Longitudinal Changes in DMP Outcome Measures

Plots of the change in the PLM from baseline for each of the 7 different DMP outcome measures are shown in Figure 2. The plots show that all measures exhibited a significant

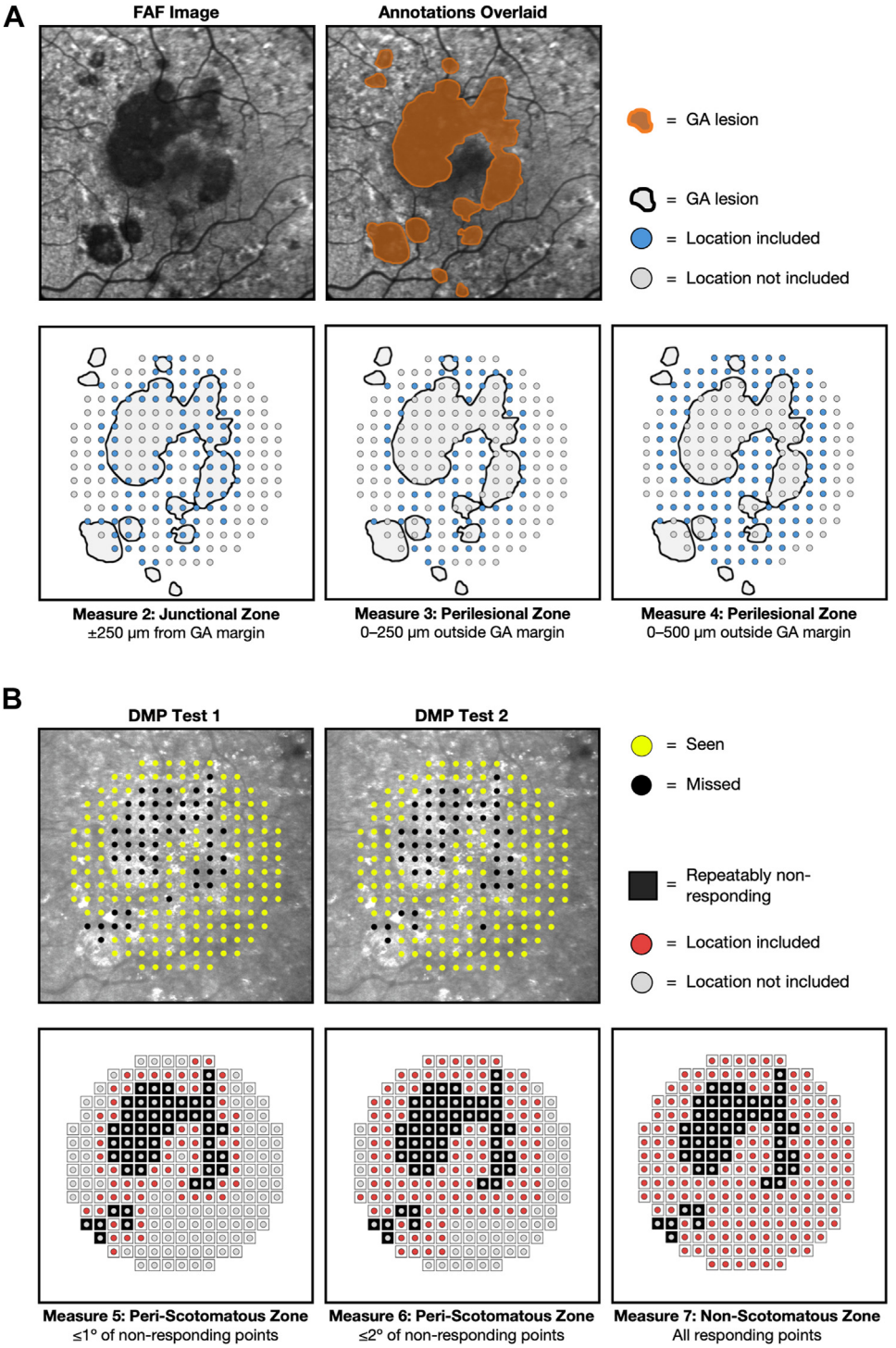

Figure 1. Illustration of how different defect-mapping microperimetry (DMP) outcome measures were derived. (**A**) Fundus autofluorescence (FAF) images were manually annotated for geographic atrophy (GA) lesion(s) and coregistered with the near-infrared fundus image obtained during DMP testing. The proportion of locations missed (PLM) on DMP were then derived from a subset of points that were within ±250 µm of the GA margin (Measure 2), between 0–250 µm outside the GA margin (Measure 3), or between 0–500 µm outside the GA margin (Measure 4). (**B**) Two DMP tests at baseline were used to identify test locations that were repeatably nonresponding (or locations that were missed on both tests). The PLM on DMP was then also derived for a subset of points that were ≤1° or 2° (along the horizontal or vertical axis) neighboring repeatably nonresponding test locations at baseline (Measures 5 and 6, respectively), or when evaluating all points excluding the repeatably nonresponding locations (Measure 7). For reference, the PLM from all test locations was also derived (Measure 1; not shown).

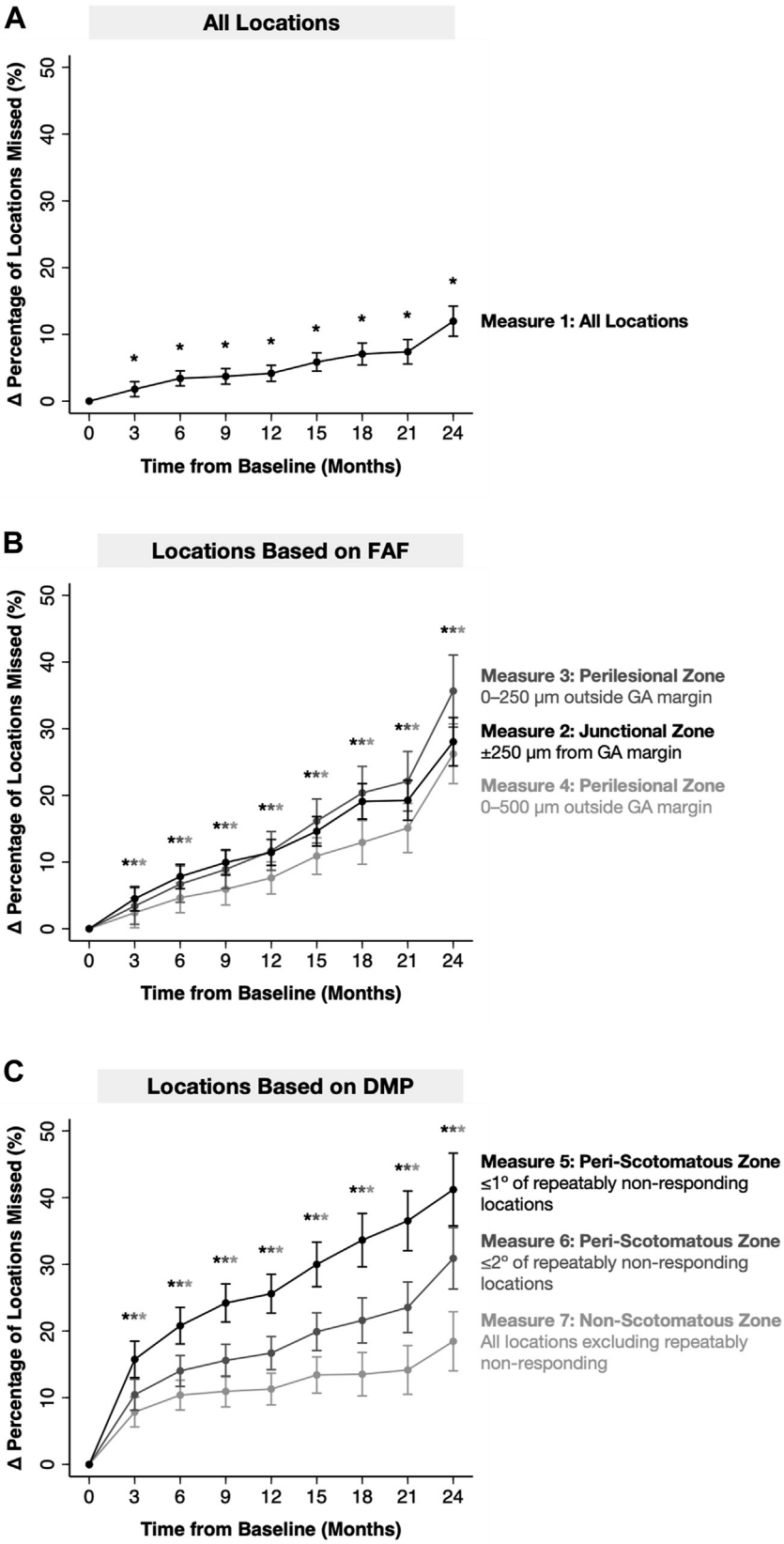

Figure 2. Plots of the change (Δ) in the proportion of locations missed (PLM) on defect-mapping microperimetry (DMP) from baseline when considering (**A**) all locations, (**B**) a subset of locations falling within the junctional or perilesional zone defined on fundus autofluorescence (FAF) imaging

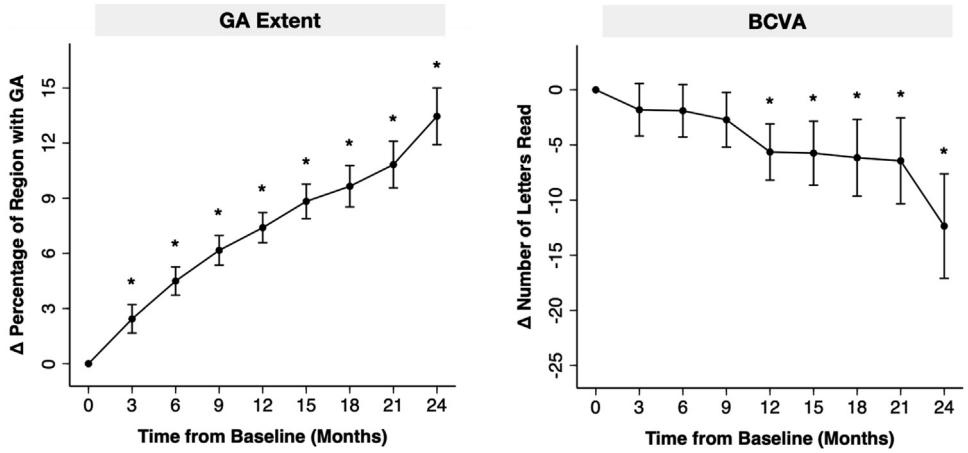

Figure 3. Plots of the change (Δ) from baseline in (left) the percentage of the corresponding central 8° region sampled on defect-mapping microperimetry with geographic atrophy (GA) and (right) best-corrected visual acuity (BCVA). A significant change from baseline (at $P \leq 0.001$) was detected at all subsequent follow-up visits after baseline for GA extent and only at the 12-month and all subsequent follow-up visits for BCVA. Asterisks indicate visits showing a significant change (at $P < 0.05$) from baseline.

increase in PLM from baseline at follow-up visits (from 3 to 24 months; $P \leq 0.046$ for all).

For comparison, plots of the change in percentage of GA in the corresponding central 8° radius region sampled on DMP testing and BCVA from baseline are shown in Figure 3. A significant increase in GA extent from baseline was also detected at all follow-up visits ($P < 0.001$ for all), but a significant decline in BCVA from baseline was only detected at the 12-month and all subsequent visits ($P \leq 0.001$ for all).

Performance for Detecting Longitudinal Changes

The CoV for all 7 DMP outcome measures are summarized and compared in Table 1. The CoV was the highest when evaluating the PLM at all test locations (Measure 1; 94%), while the CoV was the lowest when evaluating the PLM in the junctional zone on FAF (Measure 2; 59%) and peri-scotomatous zone, defined by test locations ≤1° neighboring nonresponding test locations at baseline (Measure 5; 47%). The latter 2 measures had significantly lower CoVs than the PLM at all test locations (Measure 1; $P < 0.001$ for both) and were not significantly different from the CoV for GA extent in the central 8° region (59%; $P \geq 0.097$ for both). Note that all 7 DMP measures had a significantly lower CoV than BCVA (188%; $P \leq 0.002$).

Implications for Power in GA Trials

The total sample sizes required to detect a ≥30% treatment effect in a 2-arm trial for the different DMP outcome measures are shown in Table 2. The required sample sizes over a 24-month period, when performing 2 baseline tests and follow-up testing at 6-monthly intervals, were 618, 220, and 146 participants when evaluating the PLM at all test locations (Measure 1), the junctional zone on FAF (Measure 2), and the peri-scotomatous zone defined by test locations ≤1° neighboring repeatably nonresponding test locations at baseline (Measure 5), respectively. For comparison, a total of 268 and 2428 participants would be required when evaluating the corresponding extent of GA in the central 8° region tested on DMP and BCVA, respectively. The best-performing DMP outcome measure (Measure 5) thus enabled a 46% and 94% reduction in the sample size required for GA trials over this period if this outcome measure was used.

Discussion

This study showed that customized evaluation of the PLM in a subset of points immediately adjacent (≤1°) to repeatably nonresponding locations at baseline, or at the junctional zone of GA defined on FAF imaging, outperformed the standard evaluation of all test locations on DMP for capturing longitudinal changes. Evaluating the PLM in a subset of points immediately adjacent to repeatably nonresponding locations at baseline on DMP enabled a 46% and 94% reduction in the sample size requirements in a 24-month clinical trial when compared with evaluating GA extent and BCVA, respectively. These findings demonstrate the promise of this approach for providing a much-needed, sensitive visual function outcome measure to capture the potential beneficial effects of treatments for slowing GA progression.

(with 3 different measures based on the distance relative to the geographic atrophy [GA] margin), and (**C**) a subset of locations falling in the peri-scotomatous zone (either ≤1° or ≤2° neighboring repeatably nonresponding locations from the 2 baseline DMP tests, along the horizontal or vertical meridians) or the nonscotomatous zone (all points excluding those that were repeatably nonresponding on the 2 baseline tests). Asterisks indicate visits showing a significant change (at $P < 0.05$) from baseline.

Table 1. Comparison of the Performance of the Different Measures for Detecting Longitudinal Changes on Defect-Mapping Microperimetry in Eyes with Geographic Atrophy

Outcome Measure	CoV (95% CI)	*P* Value, vs. DMP Measure		
		1	*2*	*5*
DMP: Proportion of locations missed at:				
Measure 1: All locations	94 (70–130)	-	<0.001	<0.001
Measure 2: Junctional zone on FAF, ±250 μm*	59 (45–73)	<0.001	-	0.049
Measure 3: Perilesional zone on FAF, 0–250 μm*	77 (56–94)	0.068	0.003	<0.001
Measure 4: Perilesional zone on FAF, 0–500 μm*	94 (71–116)	0.467	<0.001	<0.001
Measure 5: Peri-scotomatous zone on DMP, ≤1°†	47 (36–60)	<0.001	0.049	-
Measure 6: Peri-scotomatous zone on DMP, ≤2°†	59 (45–76)	<0.001	0.464	0.004
Measure 7: Nonscotomatous zone on DMP‡	86 (70–109)	0.340	0.009	<0.001
Percentage of GA within central 8°	58 (46–69)	<0.001	0.438	0.097
Best-corrected visual acuity	188 (134–291)	0.002	<0.001	<0.001

CI = confidence interval; CoV = coefficient of variation, defined as the standard deviation of the rate of change in the outcome measure (per year) divided by its mean (with lower values indicating better performance); DMP = defect-mapping microperimetry; FAF = fundus autofluorescence; GA = geographic atrophy.
*Relative to the GA margin based on manual annotations on FAF imaging.
†Neighboring repeatably nonresponding test locations, defined as those that were missed on both tests at baseline.
‡Defined as all test locations excluding repeatably nonresponding test locations.

A study conducted over a decade ago first reported that the absolute magnitude of change over a 24-month period in the mean sensitivity of test locations immediately adjacent to points with an absolute scotoma (from a single baseline test) was higher than the mean sensitivity of all test locations.[8] Similar results were found when evaluating the mean sensitivity of all nonscotomatous locations at baseline (term "responding" sensitivity) over time. These findings were observed when using standard threshold-based microperimetry testing with a stimulus pattern that had an interstimulus interval of 2° and a device that had a more limited dynamic range than the one used in this study.[8] It is important to note that none of these outcome measures from the previous study, which included 18 eyes from 9 individuals, were significantly associated with the change in the GA area over 24 months, except for a significant association observed with the number of absolute scotomas.[8]

However, the effectiveness of an outcome measure for capturing longitudinal changes depends on not only the magnitude of change detected but also the variability of such change over time. Using the same microperimeter, stimulus pattern, and threshold-based testing strategy as the previously mentioned previous study, a recent study compared these different microperimetry outcome measures based on their CoV—a normalized measure of performance for capturing longitudinal changes.[19] This was performed in a cohort with >20 times the number of participants who underwent microperimetry testing at 24-week intervals over a 72-week period.[19] This recent study showed that the evaluation of "perilesional" (based on locations neighboring absolute scotomas from a single baseline test) and "responding" sensitivity reduced the CoV by 30% and 20%, respectively, at 72 weeks, when compared with the evaluation of mean sensitivity from all test locations. Although the CoV of these measures was nearly half of what was observed with BCVA, it was still approximately 55% to 80% higher when compared with measurements of GA extent.[19]

Instead, we show in this study with a novel DMP testing strategy that the outcome measure of PLM from a subset of points immediately adjacent (≤1°) to repeatably nonresponding locations at baseline, or at the junctional zone of GA, had a similar CoV to the measure of GA extent in the central 8° region tested on DMP and was 69% to 75% lower than BCVA over a follow-up period of up to 24 months. The superior performance of DMP testing compared with GA extent and BCVA, relative to what was reported with standard threshold-based microperimetry testing in the recent study,[19] likely reflects the greater effectiveness of DMP testing for capturing visual sensitivity losses in eyes with GA. This is further supported by our previous findings, which showed markedly stronger structure–function correlation observed with DMP at cross section[17] and longitudinally[18] than previously reported with standard threshold-based testing.[10] Nonetheless, note that although several approaches based on evaluating a customized subset of test locations on DMP were more effective than evaluating all test locations at capturing longitudinal changes, their correlation with changes in GA extent was lower when compared with changes based on evaluating global changes on DMP testing (*data not shown*). This is expected as evaluating a customized subset of test locations on DMP does not capture progressive loss of visual sensitivity at all possible regions, but rather within regions most likely to exhibit such progression.

The CoV of evaluating the PLM as an outcome measure also increased from 47% when evaluating a subset of points within ≤1° adjacent to the repeatably nonresponding test locations at baseline to 59% when evaluating points within ≤2°. We also observed a >40% reduction in the sample size requirements for a 24-month trial when evaluating the outcome measure that included these points ≤1° to the

Table 2. Total Sample Size Estimates for Geographic Atrophy Trials Required to Detect a ≥30% Treatment Effect Using Different Defect-Mapping Microperimetry Outcome Measures in A 2-Arm Trial

Outcome Measure	Follow-Up Duration	
	12 Mos	*24 Mos*
DMP: Proportion of locations missed at:		
Measure 1: All locations	774	618
Measure 2: Junctional zone on FAF, ±250 μm*	324	220
Measure 3: Perilesional zone on FAF, 0–250 μm*	512	418
Measure 4: Perilesional zone on FAF, 0–500 μm*	748	668
Measure 5: Peri-scotomatous zone on DMP, ≤1°†	202	146
Measure 6: Peri-scotomatous zone on DMP, ≤2°†	304	254
Measure 7: Nonscotomatous zone on DMP‡	614	564
Percentage of GA within central 8°	290	268
Best-corrected visual acuity	3130	2428

Sample size estimates are based on having 80% power and a 2-sided α of 0.05, assuming 10% attrition per year, assuming 6-monthly follow-up intervals with 2 baseline visits.
DMP = defect-mapping microperimetry; FAF = fundus autofluorescence; GA = geographic atrophy.
*Relative to the GA margin based on manual annotations on FAF imaging.
†Neighboring repeatably nonresponding test locations, defined as those that were missed on both tests at baseline.
‡Defined as all test locations excluding repeatably nonresponding test locations.

repeatably nonresponding locations at baseline compared with the measure that included points ≤2°. This likely reflects how the slowly progressive nature of GA growth requires a higher spatial density of testing to better sample the regions where visual function loss is most likely to occur over time, which can be more effectively achieved with DMP compared with standard threshold-based testing.

The greater performance of detecting longitudinal changes based on evaluating points at the GA junctional zone on FAF imaging compared with all test locations in this study is also broadly supported by findings from a recent landmark trial.[11] This trial reported post hoc results of a significant beneficial treatment effect on the number of absolute scotomas at the junctional zone in one of the treatment arms,[11] while no significant treatment benefit was seen when evaluating the mean sensitivity from all locations tested, with standard threshold-based microperimetry.

This study is the first to compare different approaches for customizing the selection of test locations on DMP for longitudinal evaluation, using either the baseline functional testing results or based on anatomic changes. Slightly better performance based on the CoV was observed when evaluating the PLM in a subset of points immediately adjacent to nonresponding locations than at the junctional zone of GA at baseline, which were the best-performing outcome measures when points were selected based on baseline functional and structural data, respectively. Additionally, the functionally guided approach reduced the required sample size needed for a 24-month trial by approximately a third compared with the imaging-based approach. This is unsurprising because the former evaluates points adjacent to those directly measured as being nonresponding, which the imaging-based approach attempts to identify indirectly. Given the slightly greater complexity of the imaging-based approach (because of the need to coregister the microperimetry and annotated FAF imaging data), functionally guided selection of a customized subset of prespecified points for evaluating changes in the PLM would be the preferred approach for GA trials.

The findings of this study challenge the impression from recent landmark GA trials that detecting beneficial treatment effects on a visual function outcome is difficult.[11,12,19] Instead, this study shows that evaluating changes in the PLM from a custom-selected subset of points on DMP would be as feasible as evaluating the anatomical endpoint of GA growth and may even require fewer participants. Indeed, only approximately 200 to 150 participants in total would be needed over a 12- and 24-month period, respectively, to detect a ≥30% treatment effect in a 2-arm trial using this outcome measure. Defect-mapping microperimetry testing in this study with the stimulus pattern that sampled the central 8° radius region (4.6 mm diameter region) with a 1° interstimulus interval was also highly feasible. Each test required approximately 6 minutes to perform, having a comparable duration to a typical standard automated perimetry test. We thus recommend DMP testing and this outcome measure for future GA treatment trials.

A key limitation of this study is the small cohort size; however, the robustness of the estimates obtained in this study was achieved by the inclusion of 2 DMP tests at each visit and 3-monthly follow-up intervals for up to 24 months. Although the median baseline GA extent of the eyes included in this study was also approximately half of the mean or median extent of those included in numerous previous GA trials, the rate of GA growth in our cohort was comparable with those observed in the sham or placebo arm of these previous trials.[11,12,30–34] Therefore, the findings of this study are likely to be generalizable to future trials that use similar inclusion criteria such as these previous GA trials.[11,12,30–34] Nonetheless, further studies are needed to confirm the utility of DMP testing and the customized evaluation of its results in individuals with GA across a larger spectrum of disease severity, including variation in BCVA.

In conclusion, this study showed that evaluating longitudinal changes in the PLM on DMP at custom-selected points adjacent to nonresponding locations at baseline, or at the GA junctional zone, significantly outperformed evaluating all points. Evaluating the PLM at the points adjacent to nonresponding locations at baseline reduced the sample size needed for a 24-month trial by 46% and 94% compared with evaluating GA extent and BCVA, respectively. These findings show how this novel outcome measure could help effectively capture potential beneficial effects on visual function preservation for treatments that slow GA growth.

Footnotes and Disclosures

Originally received: January 7, 2025.
Final revision: February 24, 2025.
Accepted: March 10, 2025.
Available online: March 14, 2025. Manuscript no. XOPS-D-25-00014R1.

[1] Centre for Eye Research Australia, Royal Victorian Eye and Ear Hospital, East Melbourne, Australia.

[2] Ophthalmology, Department of Surgery, The University of Melbourne, Melbourne, Australia.

Disclosures:

All authors have completed and submitted the ICMJE disclosures form.

The authors have made the following disclosures:

R.H.G.: Consultant − Roche/Genentech, Bayer, Apellis, Belite Bio, Ocular Therapeutix, Complement Therapeutics, Boehringer Ingelheim Pharmaceuticals, Character Bioscience, Janssen, AbbVie, Astellas; Payment for expert testimony − Apellis; Support for attending meetings and travel − Bayer, Roche.

P.v.W.: Consultant − Roche/Genentech, Bayer, Novartis, Mylan.

The other authors have no proprietary or commercial interest in any materials discussed in this article.

Supported by the National Health and Medical Research Council of Australia (GNT2008382 [Z.W.] and GNT1194667 [R.H.G.]), the Bright-Focus Foundation (#M2019073 [Z.W. and R.H.G.]) and Apellis Pharmaceuticals (investigator-initiated trial grant [Z.W. and R.H.G.]). CERA receives operational infrastructure support from the Victorian Government. The funders had no role in the manuscript writing and the decision to submit the manuscript for publication.

HUMAN SUBJECTS: Human subjects were included in this study. This study was conducted in accordance with the International Conference on Harmonization Guidelines for Good Clinical Practice and with the tenets of the Declaration of Helsinki. Institutional Review Board approval was obtained for this study from the Royal Victorian Eye and Ear Hospital, and all participants provided informed consent.

No animal subjects were used in this study.

Author Contributions:

Conception and design: Guymer, Wu

Data collection: Guymer, Hodgson, Glover, Gee, Wu

Analysis and interpretation: Saeed, Guymer, Hadoux, Jannaud, Dang, Hodgson, Glover, Gee, Wu

Obtained funding: Guymer, Wu

Overall responsibility: Guymer, Wu

Abbreviations and Acronyms:
AMD = age-related macular degeneration; **BCVA** = best-corrected visual acuity; **CoV** = coefficient of variation; **dB** = decibels; **DMP** = defect-mapping microperimetry; **FAF** = fundus autofluorescence; **GA** = geographic atrophy; **IQR** = interquartile range; **PLM** = proportion of locations missed.

Keywords:
Age-related macular degeneration, Fundus autofluorescence, Geographic atrophy, Microperimetry, Visual field tests.

Correspondence:
Zhichao Wu, BAppSc(Optom), PhD, Level 7, 32 Gisborne Street, East Melbourne, VIC 3002, Australia. E-mail: wu.z@unimelb.edu.au.

References

1. Sarks JP, Sarks SH, Killingsworth MC. Evolution of geographic atrophy of the retinal pigment epithelium. *Eye (Lond)*. 1988;2:552−577.
2. Kim SY, Sadda S, Humayun MS, et al. Morphometric analysis of the macula in eyes with geographic atrophy due to age-related macular degeneration. *Retina*. 2002;22:464−470.
3. Bhutto I, Lutty G. Understanding age-related macular degeneration (AMD): relationships between the photoreceptor/retinal pigment epithelium/Bruch's membrane/choriocapillaris complex. *Mol Aspects Med*. 2012;33:295−317.
4. Bird AC, Phillips RL, Hageman GS. Geographic atrophy: a histopathological assessment. *JAMA Ophthalmol*. 2014;132:338−345.
5. Li M, Dolz-Marco R, Huisingh C, et al. Clinicopathologic correlation of geographic atrophy secondary to age-related macular degeneration. *Retina*. 2019;39:802−816.
6. Sunness JS, Rubin GS, Applegate CA, et al. Visual function abnormalities and prognosis in eyes with age-related geographic atrophy of the macula and good visual acuity. *Ophthalmology*. 1997;104:1677−1691.
7. Sunness JS, Gonzalez-Baron J, Applegate CA, et al. Enlargement of atrophy and visual acuity loss in the geographic atrophy form of age-related macular degeneration. *Ophthalmology*. 1999;106:1768−1779.
8. Meleth AD, Mettu P, Agrón E, et al. Changes in retinal sensitivity in geographic atrophy progression as measured by microperimetry. *Invest Ophthalmol Vis Sci*. 2011;52:1119−1126.
9. Varma R, Souied EH, Tufail A, et al. Maximum reading speed in patients with geographic atrophy secondary to age-related macular degeneration. *Invest Ophthalmol Vis Sci*. 2018;59:AMD195−AMD201.
10. Heier JS, Pieramici D, Chakravarthy U, et al. Visual function decline resulting from geographic atrophy: results from the chroma and spectri phase 3 trials. *Ophthalmol Retina*. 2020;4:673−688.
11. 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. *Lancet*. 2023;402:1434−1448.
12. Khanani AM, Patel SS, Staurenghi G, et al. Efficacy and safety of avacincaptad pegol in patients with geographic atrophy (GATHER2): 12-month results from a randomised, double-masked, phase 3 trial. *Lancet*. 2023;402:1449−1458.
13. Csaky K, Ferris F, Chew EY, et al. Report from the NEI/FDA endpoints workshop on age-related macular degeneration and inherited retinal diseases. *Invest Ophthalmol Vis Sci*. 2017;58:3456−3463.
14. Wickström K, Moseley J. Biomarkers and surrogate endpoints in drug development: a European regulatory view. *Invest Ophthalmol Vis Sci*. 2017;58:BIO27−BIO33.
15. FDA. Cellular, Tissue, and Gene Therapies Advisory Committee. CTGTAC Meeting #52. *Cellular and Gene Therapies for Retinal Disorders*. 2011:2016. v.
16. Wu Z, Cimetta R, Caruso E, Guymer RH. Performance of a defect-mapping microperimetry approach for characterizing

progressive changes in deep scotomas. *Transl Vis Sci Technol*. 2019;8:16.
17. Wu Z, Hadoux X, Jannaud M, et al. Visual sensitivity loss in geographic atrophy: structure−function evaluation using defect-mapping microperimetry. *Invest Ophthalmol Vis Sci*. 2024;65:36.
18. Saeed A, Guymer RH, Hadoux X, et al. Longitudinal changes of visual sensitivity with geographic atrophy progression assessed by defect-mapping microperimetry. *Invest Ophthalmol Vis Sci*. 2024;65:13.
19. Chang DS, Callaway NF, Steffen V, et al. Macular sensitivity endpoints in geographic atrophy: exploratory analysis of chroma and spectri clinical trials. *Ophthalmol Sci*. 2024;4:100351.
20. Velaga SB, Nittala MG, Hariri A, Sadda SR. Correlation between fundus autofluorescence and en face OCT measurements of geographic atrophy. *Ophthalmol Retina*. 2022;6:676−683.
21. Wu Z, Ayton LN, Guymer RH, Luu CD. Intrasession test−retest variability of microperimetry in age-related macular degeneration. *Invest Ophthalmol Vis Sci*. 2013;54:7378−7385.
22. Wu Z, Hadoux X, Jannaud M, et al. Systematic underestimation of visual sensitivity loss on microperimetry: implications for testing protocols in clinical trials. *Transl Vis Sci Technol*. 2023;12:11.
23. 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:BIO158−BIO167.
24. Noyel G, Thomas R, Bhakta G, et al. Superimposition of eye fundus images for longitudinal analysis from large public health databases. *Biomed Phys Eng Express*. 2017;3:45015.
25. Spaide T, Jiang J, Patil J, et al. Geographic atrophy segmentation using multimodal deep learning. *Transl Vis Sci Technol*. 2023;12:10.
26. Goh KL, Abbott CJ, Hadoux X, et al. Hyporeflective cores within drusen: association with progression of age-related macular degeneration and impact on visual sensitivity. *Ophthalmol Retina*. 2022;6:284−290.
27. Goh KL, Kumar H, Hadoux X, et al. Hyperpigmentary abnormalities in age-related macular degeneration: association with progression and impact on visual sensitivity. *Br J Ophthalmol*. 2024;108:263−267.
28. Cheng QE, Gao J, Kim BJ, Ying G-S. Design characteristics of geographic atrophy treatment trials: systematic review of registered trials in ClinicalTrials.gov. *Ophthalmol Retina*. 2018;2:518−525.
29. Iddi S, Donohue MC. Power and sample size for longitudinal models in R − the longpower Package and Shiny app. *R J*. 2022;14:264−282.
30. 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;136:666−677.
31. Rosenfeld PJ, Dugel PU, Holz FG, et al. Emixustat hydrochloride for geographic atrophy secondary to age-related macular degeneration: a randomized clinical trial. *Ophthalmology*. 2018;125:1556−1567.
32. Liao DS, Grossi FV, El Mehdi D, et al. Complement C3 inhibitor pegcetacoplan for geographic atrophy secondary to age-related macular degeneration: a randomized phase 2 trial. *Ophthalmology*. 2020;127:186−195.
33. Jaffe GJ, Westby K, Csaky KG, et al. C5 inhibitor avacincaptad pegol for geographic atrophy due to age-related macular degeneration: a randomized pivotal phase 2/3 trial. *Ophthalmology*. 2021;128:576−586.
34. Rosenfeld PJ, Berger B, Reichel E, et al. A randomized phase 2 study of an anti−amyloid β monoclonal antibody in geographic atrophy secondary to age-related macular degeneration. *Ophthalmol Retina*. 2018;2:1028−1040.