

Model Structure Uncertainty in the Characterization and Growth of Geographic Atrophy

Janan Arslan^{1,2}, Kurt K. Benke^{3,4}, Gihan Samarasinghe⁵, Arcot Sowmya⁵,
Robyn H. Guymer^{1,2}, and Paul N. Baird²

¹ Centre for Eye Research Australia, University of Melbourne, Royal Victorian Eye & Ear Hospital, East Melbourne, Victoria, Australia

² Department of Surgery, Ophthalmology, University of Melbourne, Parkville, Victoria, Australia

³ School of Engineering, University of Melbourne, Parkville, Victoria, Australia

⁴ Centre for AgriBioscience, AgriBio, Bundoora, Victoria, Australia

⁵ School of Computer Science and Engineering, University of New South Wales, Kensington, NSW, Australia

Correspondence: Janan Arslan, Centre for Eye Research Australia Ltd, Level 7, 32 Gisborne St., East Melbourne, VIC 3002, Australia. e-mail: janan.arslan@unimelb.edu.au

Received: August 6, 2020

Accepted: March 10, 2021

Published: May 3, 2021

Keywords: age-related macular degeneration; geographic atrophy; linear model; model structure uncertainty

Citation: Arslan J, Benke KK, Samarasinghe G, Sowmya A, Guymer RH, Baird PN. Model structure uncertainty in the characterization and growth of geographic atrophy. *Transl Vis Sci Technol.* 2021;10(6):2. <https://doi.org/10.1167/tvst.10.6.2>

Purpose: To identify the most suitable model for assessing the rate of growth of total geographic atrophy (GA) by analysis of model structure uncertainty.

Methods: Model structure uncertainty refers to unexplained variability arising from the choice of mathematical model and represents an example of epistemic uncertainty. In this study, we quantified this uncertainty to help identify a model most representative of GA progression. Fundus autofluorescence (FAF) images and GA progression data (i.e., total GA area estimation at each presentation) were acquired using Spectralis HRA+OCT instrumentation and RegionFinder software. Six regression models were evaluated. Models were compared using various statistical tests, [i.e., coefficient of determination (r^2), uncertainty metric (U), and test of significance for the correlation coefficient, r], as well as adherence to expected physical and clinical assumptions of GA growth.

Results: Analysis was carried out for 81 GA-affected eyes, 531 FAF images (range: 3–17 images per eye), over median of 57 months (IQR: 42, 74), with a mean baseline lesion size of $2.62 \pm 4.49 \text{ mm}^2$ (range: 0.11–20.69 mm^2). The linear model proved to be the most representative of total GA growth, with lowest average uncertainty (original scale: $U = 0.025$, square root scale: $U = 0.014$), high average r^2 (original scale: 0.92, square root scale: 0.93), and applicability of the model was supported by a high correlation coefficient, r , with statistical significance ($P = 0.01$).

Conclusions: Statistical analysis of uncertainty suggests that the linear model provides an effective and practical representation of the rate and progression of total GA growth based on data from patient presentations in clinical settings.

Translational Relevance: Identification of correct model structure to characterize rate of growth of total GA in the retina using FAF images provides an objective metric for comparing interventions and charting GA progression in clinical presentations.

Introduction

Geographic atrophy (GA) is one of the two late stages of the debilitating eye disease age-related macular degeneration (AMD). The current estimate for the number of affected patients globally is approximately 5 million, with the number of GA cases expected to increase to 9 to 10 million patients by 2040.¹ The etiology of GA remains elusive, and no drug

therapies are available.² GA is a progressive disease, and vision will continue to deteriorate with a possibility of legal blindness. The deterioration in vision is associated with the growth of GA lesions. GA is defined as dead retinal pigment epithelium (RPE) and photoreceptor cells with closure of the underlying choriocapillaris.^{2,3} They appear as sharply demarcated areas, which are traditionally identified by retinal imaging.² The appearance of GA lesions in the macular region affects vision, and the severity of vision loss is linked

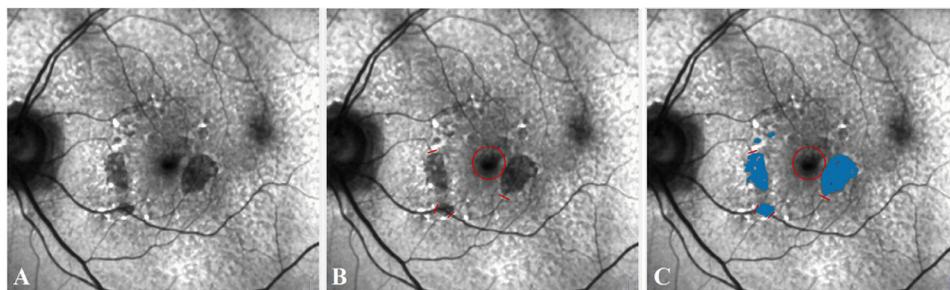


Figure 1. RegionFinder annotation process. (A) Original image, (B) fovea and blood vessels restricted, and (C) final annotation with region-growing algorithm. Prior to annotation, the end-user must restrict nonlesion areas such as the fovea, optic disc, and blood vessels. This preventative measure ensures the region-growing annotation does not “spill” into features similar in color intensity to that of lesions. The lesions can then be annotated using the mouse cursor.

with the size and the location of the lesions in the macula.^{4,5} The pattern of GA growth is not well understood, and research publications to-date have often described the trend in GA growth in qualitative terms rather than quantitatively. It would be useful to have an objective and quantitative metric for GA progression as a means of trend identification and to predict the rate of growth of GA. A predictive model would inform the patient and the clinician about disease progress and inform on the validity of any future interventions.

Past studies of GA growth are often based on local growth of a single lesion, with “effective” radius plotted against time (i.e., radial growth), rather than growth of global or total GA.^{1,6} Past “qualitative” observations have suggested a possible linear progression of GA lesion area.^{1,5,7} Dreyhaupt et al.⁷ modeled the natural course of GA using linear and exponential mixed-effects models fitted to areas of atrophy computed from fundus autofluorescence (FAF) images.⁷ They found that the linear and exponential models were similar in performance. The Age-Related Eye Disease Study (AREDS) Report Number 26 suggested that, at least for fundus photographs, a linear model growth was superior to a quadratic model for different lesion sizes. The AREDS study report suggested that a linear relationship may in part be related to overlapping areas of atrophy (rather than the expansion of a single lesion over time).⁸ Despite anecdotal observations for individual lesions, there appears to be very little reported on modeling total GA progression (i.e., global growth of all lesions combined). The variability in many study findings has been attributed to errors associated with the accuracy and precision of assessment methods and therefore epistemic uncertainty also needs to be addressed in model development.⁶

Epistemic uncertainty is due to lack of knowledge and refers to reducible errors, such as subjective

uncertainty, measurement error, data entry error, or using the wrong model.^{9–11} In the case of GA assessment, the most relevant epistemic uncertainty is “model structure uncertainty” (i.e., identifying the correct model for disease progression). This uncertainty can be due to limited availability of comprehensive datasets, incomplete knowledge of the disease etiology and pathogenesis, and errors from measuring equipment that can lead to the presence of imprecise and uncertain data (i.e., measurement error and noisy data). This makes the process of developing an appropriate model challenging. For example, although the RegionFinder software provides a fast, consistent, and semiautomated process for area segmentation, it relies on human-user input for its function. A grader relies on judgement and experience when annotating GA lesions (Fig. 1). Methods available to assess structural problems include model checking (e.g., goodness-of-fit tests, calibration test, residual error assessment) and comparing tested predictions against independent data.^{12,13} There appear to be no publications that have quantified epistemic uncertainties in GA measurement. Also, the authors are unaware of any publications that have incorporated uncertainty analysis when modeling the progression of GA.

The objective of this study was to address model structure uncertainty (an epistemic uncertainty) by investigating a suitable model for characterizing global GA progression in FAF images in the context of routine clinical presentations for monitoring GA progression. The analysis required (1) a process for statistical model selection, (2) application of an uncertainty metric, (3) graphic evaluation to show that growth patterns meet physical and clinical assumptions consistent with current understanding of GA progression, and (4) an objective metric for quantifying the rate of GA progression. Six statistical regression models were evaluated using time-series FAF images of GA to

find the most appropriate GA growth model for trend analysis.

Methods

Data

The study was approved by the Human Research Ethics Committee of the Royal Victorian Eye and Ear Hospital (RVEEH). The study was conducted in accordance with the International Conference on Harmonization Guidelines for Good Clinical Practice and tenets of the Declaration of Helsinki. Ethics approval was provided by the Human Research Ethics Committee (HREC: Project No. 95/283H/15) by the RVEEH. Written informed consent was obtained from all participants.

Data for GA-affected eyes were collected retrospectively from the Centre for Eye Research Australia Ltd (CERA) and a private ophthalmology clinic. Data collection involved time-series FAF images from a database of patient presentations in clinical settings, and GA area estimation by the Spectralis HRA+OCT instrumentation and the supplied RegionFinder algorithm for area segmentation (Heidelberg Engineering, Heidelberg, Germany). Inclusion criteria included being older than the age of 50 years, having a diagnosis of GA in either one of both eyes, and having foveal-centered FAF images from at least three clinical review visits and collected over a minimum of 2 years. An optical coherence tomography (OCT) image was taken at every visit, and the FAF image was centered by the patient viewing a reference point at the center of the image. Participants with neovascular AMD (nAMD) were excluded, as were other diseases that cause atrophy similar to GA, based on the clinical appearance determined by a retinal specialist (RHG), as well as clinical history and multimodal imaging with angiography performed to identify if there was any indication that nAMD might be present. Given that progression over time was the objective, no minimum lesion size was specified. Participants had no prior history of treatment or treatment trial for their condition. Records were anonymized to respect patient privacy.

Instrumentation

FAF is a recognized tool used to measure GA size and growth rate in longitudinal experiments requiring a high degree of reproducibility.¹⁴ First described by Delori et al.,¹⁵ the FAF (excitation wavelength 488 nm, emission > 500 nm) is an ophthalmic imaging technique

designed to capture GA as hypoautofluorescent (black) areas with sharp borders that delineate GA lesions, and hyperautofluorescent areas (bright areas that show the main fluorophore lipofuscin, but other fluorophores exist as well) that provide insight into the health and functionality of the RPE.^{15–18} The FAF images were acquired on the Spectralis HRA+OCT instrument. Pupils were dilated at every visit before image acquisition. FAF images with 30° × 30° field of view were captured. The dataset consisted of images in both high speed (768 × 768 pixels), and high resolution (1536 × 1536 pixels) formats. The Automatic Real-Time Tracking (ART) was also recorded for each image.

Estimation of GA Area

The image analysis for area estimation was performed using the semiautomated software algorithm RegionFinder for image segmentation of lesions. The software enables graders to annotate GA lesions, extract measurements of lesions, and records changes in longitudinal time-series images.^{19–22} Published evaluations confirm that the RegionFinder algorithm is accurate, reproducible, and time-efficient for identification and quantification of lesions.^{19,21,22} It is currently incorporated into the Spectralis instrument and used in clinical practice.

Square Root Transformation

In the past, application of the square root transformation of GA area for assessment of GA growth rate has been associated with a number of benefits, such as (1) it reduces test–retest variability, (2) it creates uniformity in intergrader differences across a range of lesion sizes, and (3) it reduces baseline GA size dependency from the GA growth rate. Furthermore, the square root transformation is not affected by zeroes and extremely small values.

Yehoshua et al.²³ investigated the reproducibility of GA area measurements and enlargement rate of GA, including usefulness of the square root transformation. They found that it eliminated GA baseline size dependency from the GA growth rate. The correlation between lesion size and test–retest standard deviations was significant with respect to original GA area (Pearson's $r = 0.60$, $P < 0.001$; Spearman's $\rho = 0.73$, $P < 0.001$). However, when a square root transformation of the lesion area measurements was performed prior to test–retest standard deviation calculations, the correlation between baseline lesion size and test–retest standard deviations was no longer apparent (Pearson's $r = 0.07$, $P = 0.72$; Spearman's $\rho = 0.12$, $P = 0.51$).

Pfau et al.²⁴ quantified lesion progression using a linear mixed-effects model with two-level random effects (i.e., eye- and patient-specific effects) and shape-descriptive factors. The authors normalized the variables for the lesion area, perimeter, and circularity using the square root transformation.²⁴ Monés and Biarnés²⁵ assessed the progression of GA and its baseline using the square root transformation, with both Pearson's r and Spearman's ρ in the assessment. They plotted the relationships using linear regression with locally weighted scatterplot smoothing curves; one plot compared GA area growth (mm^2/year) against baseline GA area (mm^2), whereas the other plotted radial growth (mm/year) against the square root transformed baseline. They found the correlation between radial growth and square root-transformed baseline area was negative (Pearson's $r = -0.30$, $P = 0.0005$; Spearman's $\rho = -0.25$, $P = 0.0042$), which suggests that as lesions grow larger, the progression rate starts decreasing.²⁵ Domalpally et al.²⁶ studied a parameter—the Geographic Atrophy Circularity Index (GACI)—in the assessment of GA progression. They used regression analysis to assess the relationship between baseline characteristics and annual progression rates of GA. Similar to Monés and Biarnés,²⁵ they found statistically significant correlations between GACI and growth rate in mm^2 ($r = -0.31$, $P < 0.001$) and GACI and square root-transformed measurements ($r = -0.39$, $P < 0.001$).

In this study, we investigated six regression models against two outcomes: (1) the original scaled, untransformed total GA area (with a unit of mm^2/year); and (2) the square root-transformed GA area (with a unit of mm/year). We used the following criteria to determine the strength of the square root transformation for the cohort: (1) whether the square root transformation normalized the distribution of residuals from the regression models tested; (2) if the transformation linearized the growth rates, as would be expected from this type of transformation; and (3) if the transformation significantly improved the fit of the model as compared with its untransformed, original-scaled counterpart. It is important to note that, in regression analysis, the assumption of normality applies to the residuals only. The distribution of independent and dependent variables can be skewed if the residuals of the regression model are normally distributed.

Modeling and Statistical Analysis

Total GA area in the image was computed for each eye and was recorded as a time-series. Total GA area was used as one output. The second output for

comparison was the square root transformed results for GA area. We calibrated 6×2 regression models for each patient (i.e., 6 for original-scaled GA area, and 6 for square root-transformed area). These regression models were linear, logarithmic, exponential, power, a standard quadratic (Q1), and quadratic without a linear term (Q2) (refer to Appendix).

The models were assessed as follows: (1) graphical evaluation of whether the model gradient is consistent with physical and clinical assumptions of growth (i.e., comparing all models on a graph for each eye); (2) computing the coefficient of determination, r^2 , for each model for each eye (where r is the correlation coefficient), and then placing the model r^2 values into a matrix (model type \times patient ID) for comparison; (3) quantifying the uncertainty metric, U , for each model; (4) when models had fractional differences in terms of trends, r^2 and U , we compared the models using both Spearman's ρ and Pearson's r rank correlation coefficients to see if the growth trends between the models were similar or significantly different; (5) the model that satisfied steps 1–3 was selected as the best fitting model, based on the principle of *Occam's razor* (i.e., the most simple model for explaining the results is the preferred model); (6) we used the test of significance for the correlation coefficient, r , to test whether sample data were sufficient to model the relationship in this cohort correctly for every model tested for every subject; and (7) for the best fit model, we compared the normality of residuals of the model with the original GA area scale (i.e., mm^2/year) to that of the square root transformed GA area scale (i.e., mm/year) to see if they satisfied the assumptions of normality for the residuals.

The graphical interpretation demonstrated how well each model fitted the data and whether the fits were in-line with anticipated clinical progression (e.g., GA lesions will continue to grow over time). Models that did not follow the physical and clinical assumptions of growth were eliminated from further consideration. These assumptions included (1) the model should illustrate continued GA growth, not GA regression, irrespective of the speed of progress; and (2) noting that because of the limited space within the retina, the growth has an upper limit.

Once a model met the required assumptions, the r^2 values for all models were computed. These values were placed in the r^2 matrix for comparison. For example, in the linear model, r^2 is the proportion of the variance in the dependent variable, Y , that is predictable from the independent variable, X .

For each patient, the best model was chosen based on the highest r^2 value. The frequency of the best model occurrence was quantified and tabulated. For the best model chosen for each patient, the uncertainty

metric U (Appendix, Eq. 10) was calculated. Note that r^2 is the variability accounted for by the regression model itself, whereas the value U for uncertainty is the variability not accounted for by the regression model.

The average uncertainties for each model type were pooled to determine the model type that had the lowest uncertainty on average. The model with high average r^2 and low average U , that met clinical assumptions of progression, was deemed the most practical model for GA progression. We then assessed the residuals of the selected regression model (both with and without square root transformation) to see how the residuals of the model behaved and whether transformation improved the normality of residuals.

A statistical test of significance for the correlation coefficient, r , can be used to determine whether r^2 and U for each model fitted is significant, where the test is adjusted for sample size (see Appendix). In addition to this, with respect to the population of GA area measurements, the Central Limit Theorem and sample size calculations can be used to support sufficiency of sample size (i.e., the sampling distribution of the mean values tends toward a normal distribution, for $n > 30$).

Results

Segmentation Data

A total of 81 eyes from 45 patients were included in the study. A total of 36 patients had bilateral GA and 9 patients had unilateral GA. From these 81 eyes, a time-series of clinical presentations resulted in 531 FAF images (range: 3–17 images per eye, median: 6 [interquartile range {IQR}: 4, 7]) were used for GA segmentation by RegionFinder and subsequent analysis. The length of time in which patients were monitored was a median of 57 months (IQR: 42, 74). The mean baseline GA area was $2.62 \pm 4.49 \text{ mm}^2$ (range: 0.11–20.69 mm^2). The average change from last GA measurement to baseline GA area was $6.07 \pm 4.99 \text{ mm}^2$. The mean age at baseline was 76.84 ± 8.60 years (range: 55–96 years). The cohort was generally a diverse representation of GA progression, with a mixture of slow and fast progressors. The median ART was 35 (IQR: 25, 75; range: 5–100).

Graphical Data Modeling

Illustrative examples of regression models fitted to GA time-series data are depicted in Figures 2A and 2B, and Figure 3. In Figure 2A, the central line is the linear approximation, which is overlaid with the

logarithmic and Q2 models—showing minimal differences. The curved plot represents power, exponential, and Q1 quadratic model (with the Q1 fit requiring an extra parameter, limiting its use for small datasets), with the curvature attempting to accommodate outliers and violating clinical assumptions on GA growth. The correlation coefficients suggest that the linear approximation is the most effective and minimalist model. A similar pattern is evident over the full set of 81 eyes tested. In Figure 2B, another example shows that the linear approximation is simple and suited to estimation with small datasets, matching other two-parameter nonlinear regression models. In Figure 3, the limitations of the Q1 model for both interpolation and extrapolation are exhibited, even with the extra degree of freedom for fitting. The growth pattern, in a “J” curve, show both decreasing then increasing patterns, which violates the clinical assumptions of GA growth. The Q1 model was eliminated from further consideration as it did not accurately represent the expected pattern of growth.

A square root transformation is expected to improve results for the growth curve by linearizing the plot. In Figure 4, we applied the square root transformation to the “J” curve in Figure 3. The square root transformation showed a flattening effect, as expected, producing a slight improvement, but did not linearize the trajectory of the Q1 curve. This was also evident in some other cases (see following section). The trajectories for the linear and logarithmic models remained similar, with and without square root transformation.

Model Ranking by r^2

Following the elimination of Q1, the five remaining models (i.e., linear, exponential, power, logarithmic, and Q2) were evaluated by comparison of results in the r^2 matrix. That is, the models were assessed with and without transformation, with dimensions of the matrix being 2 (outcomes) \times 5 (models) \times 81 (eyes) with $810 \times r^2$ values. The average r^2 values for each eye were calculated for the five models (Table 1). The linear, logarithmic, and Q2 models were the best candidates for GA growth modeling based on average r^2 . Further examination of the logarithmic, linear, and Q2 models graphically revealed that the gradients were indistinguishable. To illustrate further, consider the case illustrated in Figure 2B. For this case, the linear and logarithmic models both had a coefficient of determination of $r^2 = 0.9995$, whereas the Q2 model had $r^2 = 0.9994$. A paired t -test of comparison for the 81 eyes showed that there was no statistically significant difference between the linear

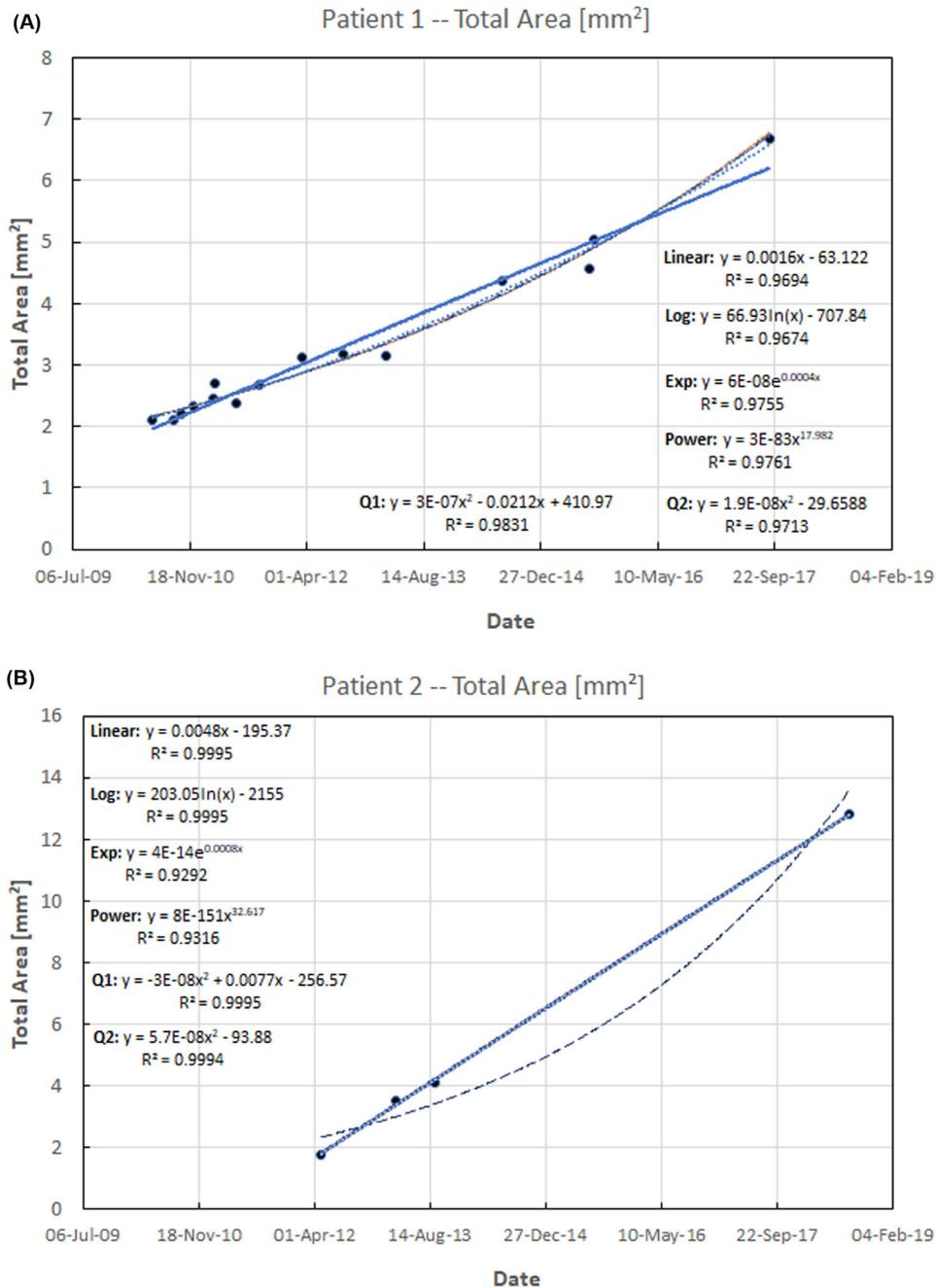


Figure 2. (A) Example of six regression models fitted to GA data for a single patient. The central line is a linear approximation overlaid with logarithmic and quadratic without linear term (Q2) models, showing minimal differences. Curved plot includes the exponential, power, and quadratic model (Q1) with concave upward curvature. (B) Another example shows how the linear model matches other overlaid models. It is well suited to estimation with small datasets and for extrapolation.

and logarithmic models (level of significance: $P < 0.05$). The correlation coefficients, Spearman's ρ and Pearson's r , were used to compare the slopes of the models, resulting in $\rho = 1$ and $r = 0.9999$. There were small differences in the coefficient of determination and the patterns of progression were similar. The linear model was deemed preferable, as it was similar to the

logarithmic and Q2 models with respect to average r^2 , and it displayed the lowest average U . Implementation and interpretation of the linear model is simple, and the linear gradient is a direct measure of rate of GA progression. The parameters of the logarithmic and Q2 models do not have straightforward interpretations.

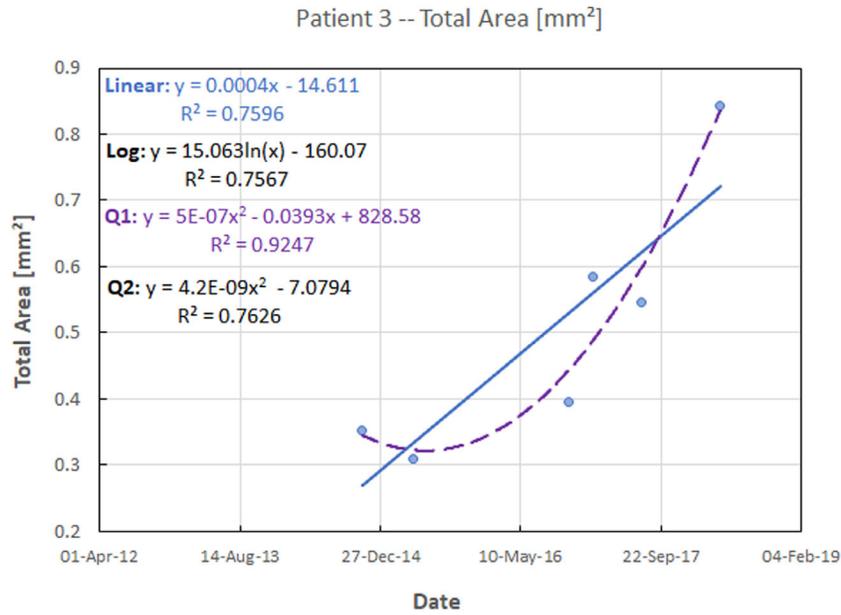


Figure 3. Illustrative example of anomalous fitting with the Q1 quadratic model. Even with three parameters, for more accurate fitting, the trendline is biologically implausible and violates assumptions on GA growth. The linear, logarithmic, and quadratic without linear term (Q2) model trendlines are indistinguishable.

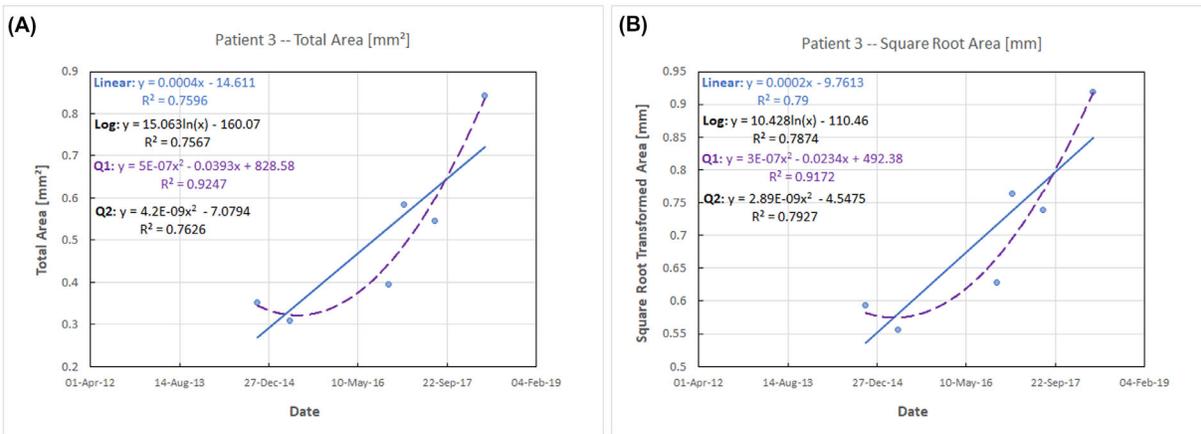


Figure 4. (A) (i.e., left image) and (B) (i.e., right image) directly in the image (i.e., similar to Fig. 2). Comparison between regression models with and without the square root transformation: (A) original scaled GA area plots, and (B) square root transformed GA area plots. The square root transformation tends to flatten but not linearize all nonlinear plots. Patterns of progression were similar in this comparison.

Quantifying Uncertainty

The metric U shows more explicitly the unexplained variability associated with the regression model, that is, uncertainty that is the sum of statistical variability and model structure uncertainty. Quantification and ranking of average uncertainty for each regression model tested revealed the linear model had the lowest average U value, with the lowest average uncertainty of 0.025 (Table 2).

Sample Size

The hypothesis test for the significance of the correlation coefficients, r , showed statistical significance ($P = 0.01$) for the GA original scale and the square root-transformed scale, providing support for the tabulated regression models. The problem of small sample size often encountered in ophthalmology has been a continuing issue because of the wide range of eye conditions that may often be rare in occurrence. Studies have

Table 1. Average r^2 for GA Area Growth Models

Model	Original Scale Average r^2	Square-Root Transformed Average r^2
Linear	0.9205	0.9299
Logarithmic	0.9196	0.9302
Quadratic (Q2)	0.9213	0.9295
Power	0.9048	0.9043
Exponential	0.9036	0.9033

Table 2. Average U for GA Area Growth Models

Model	Original Scale Average Uncertainty U	Square-Root Transformed Average Uncertainty U
Linear	0.025	0.014
Logarithmic	0.054	0.054
Quadratic (Q2)	0.028	0.026
Power	0.061	0.091
Exponential	0.135	0.135

shown that the length of the statistical confidence interval for a regression model outcome varies inversely with sample size, with the relationship flattening out for $n > 30$.^{27,28} For $n > 30$, the sampling distribution of the mean values tends toward a normal distribution according to the Central Limit Theorem. In the current study, there were $n = 81$ eyes, each having its own time-series plots of GA data, for a total of $n = 531$ for GA data, stratified so that the individual models are fitted separately to each eye.

Evaluation of Residuals

We compared the residuals of linear models using the original GA scale to that of linear models using a square root-transformed GA area. This can be investigated using a Q-Q plot (in which residuals are normally distributed, with a good-fit model having residuals closely following a 45° line, see Fig. 5), and also a histogram of residuals (i.e., checking variance and symmetry for assumption of a normal distribution). An illustrative case for the linear model is provided for patient 4 in Figure 5. It is apparent that using the original scale produces similar residuals when compared with the square root transformation.

Discussion

General Comments

The lack of knowledge regarding etiology, pathogenesis, and pattern of GA growth explains in part the paucity in treatment options. Retrospective evalu-

ation of growth over consecutive patient presentations is possible using the Spectralis HRA+OCT instrumentation and associated RegionFinder segmentation software. A missing factor is the lack of an effective model to characterize the trend in GA growth that would provide an objective metric for progression that can be used to evaluate interventions and identify fast and slow progressors.

Uncertainties affecting GA-related measurements and clinical results are statistical variability and epistemic uncertainty—defined as the uncertainties due to incomplete knowledge of the system or process. Epistemic uncertainties are reducible because identifying process limitations can lead to error minimization (e.g., upgrading instrumentation or improving operator training).²⁹ This study investigated the epistemic uncertainty referred to as “model structure uncertainty” associated with modeling GA trend and growth in patient data.

Constraints on model selection include the need for model output to be monotonically increasing with time, and to be physically plausible. Any model fitted by regression methods must pass tests including statistical significance. Six regression models were tested for characterization of the trend using the record of patient presentations. The models were linear, exponential, power, logarithmic, Q1 (original quadratic), and Q2 (quadratic without the linear term) in structure (see Appendix). All were characterized by two parameters, except for the Q1 model, which had three parameters and therefore an extra degree of freedom. Among these, the linear model was found to be the preferable model for GA progression based on simplicity and the selection criteria.

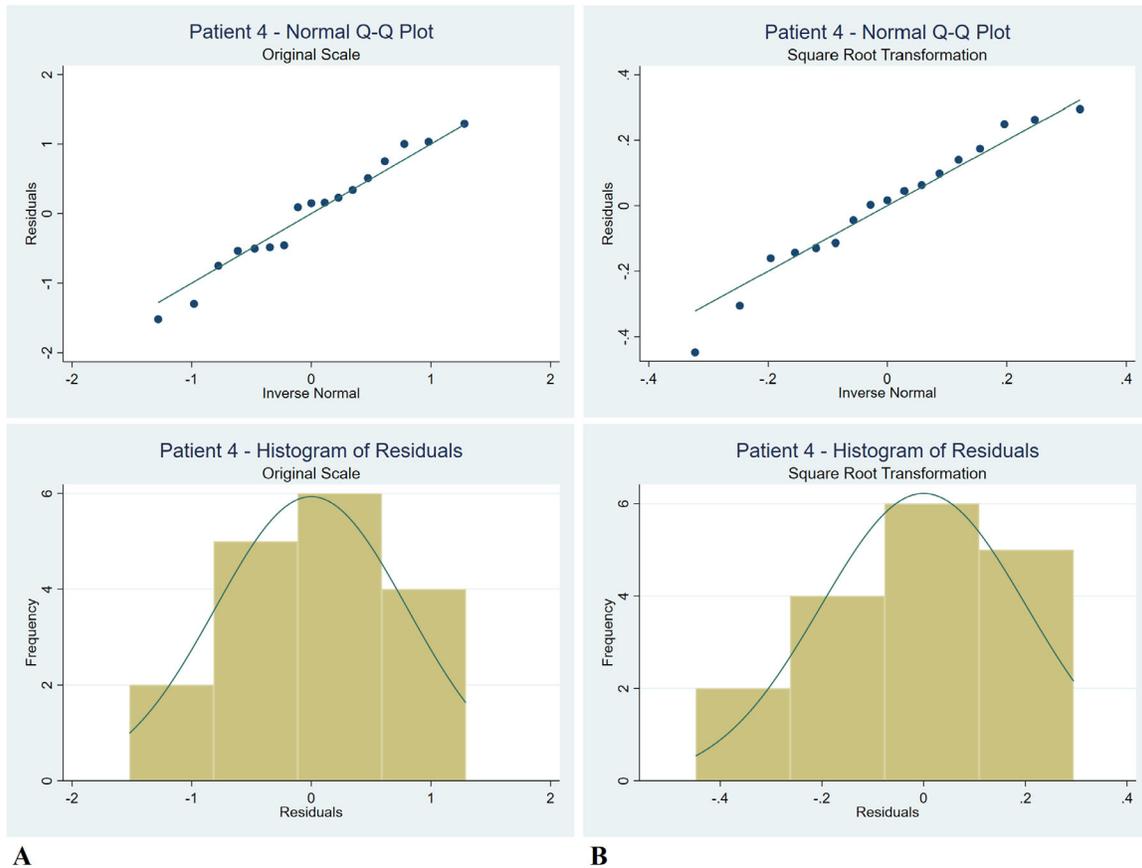


Figure 5. A Q-Q plot of residuals for the linear model for (A) original scaled GA area, and (B) square root transformed GA area. As evidenced by clustering about the 45° line, the assumption of normality in residuals appears to hold in both cases.

The linear progression model had a high average r^2 (0.9205) over 81 eyes, and the lowest average uncertainty ($U = 0.025$). These findings were consistent, even when the total GA area outcome was square root transformed ($r^2 = 0.9299$, $U = 0.014$). The models that displayed a nearly identical trend to the linear model were the logarithmic and Q2 models. When superimposing the trendlines, it was difficult to distinguish one from the other. A marginal difference was found between the average r^2 produced for linear, logarithmic, and Q2 models. However, not only did the logarithmic and Q2 models have higher average uncertainty, when compared with the linear model, there were problems associated with interpretation and application.

The Q1 model can produce a high coefficient of determination because of the extra parameter relative to the other two-parameter models. However, the concave upward growth (with trend initially decreasing to a minimum for a time then increasing again later) violated physical assumptions on the nature of GA growth over time (Fig. 3). The remaining models, exponential and power law, performed poorly, having highest average uncertainties and

graphically poor fitting around the progression data points.

The linear model appeared to be best suited for GA progression because of the advantages of (1) lowest average uncertainty, (2) realistic representation of the GA growth observed over the limited number of patient presentations, (3) ease in interpretation and extrapolation, (4) simple implementation, (5) applicability to very small datasets, and (6) simple parameters for providing rate of progression and onset estimates.

The Linear Model

Based on the analysis of patient data from the history of clinical presentations, the most statistically significant model for the rate of growth of GA was found to be the linear approximation, which is expressed more formally in the following representation:

$$A = gt + t_0$$

where A is the total area of the hypoautofluorescence as derived from the RegionFinder segmentation

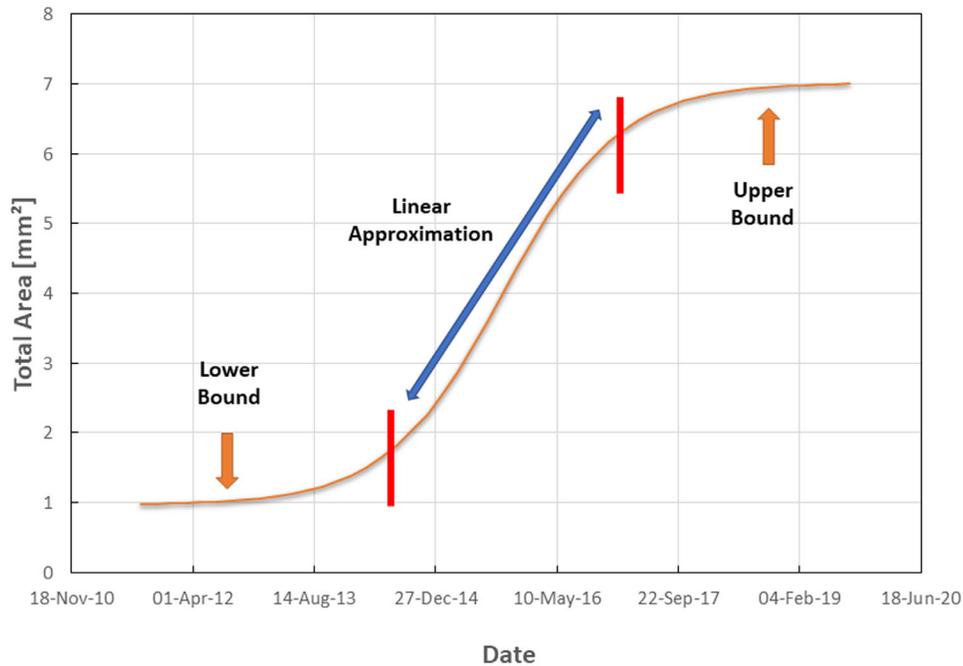


Figure 6. A linear regression is consistent with clinical assumptions of GA progression in the sense that it provides an estimate of the gradient in a growth model. The gradient is an objective metric for monitoring and comparing interventions involving the typically limited number of clinical presentations by patients.

algorithm, t is the date/time of the patient presentation, g is the gradient representing the rate of growth of GA, and t_0 is the upper bound on the date of onset. The key parameters g and t_0 are obtained from regression analysis of the history of patient presentations. The gradient parameter is an easily understood metric for GA growth for patient presentations in a clinical setting. It can be computed for small datasets and is readily incorporated into software for onscreen analytics of data recorded by the instrumentation.

The key parameter is the gradient, g , which quantifies the rate of GA progression. The value of g is an objective metric that quantifies the impact of the clinical intervention. A high value of g identifies a fast progressor, and a low value of g identifies a slow progressor. An intervention will change the gradient of the line from the time it is applied (requiring a second linear regression model from the start of the intervention). The gradient is therefore a *figure of merit* for response to the intervention.

Previous qualitative observations of GA growth have resulted in speculation that, in the case of multiple lesions, expansion in growth leads to overlapping areas, which slows down the growth of total GA.^{7,8} Although this proposed mechanism weakens the case for a Q1 model for growth in comparison to the alternative linear model, even a linear approximation would eventually be challenged by the finite area of the retina

or limited field-of-view of imaging techniques. From a clinician's perspective, total GA cannot grow indefinitely, as the size of the retina provides a physical limit in area. The assumption is that eventually the measured growth rate will stop due to termination of patient presentations as the result of patient mortality, or no further change in visual acuity tests, or it will taper-off significantly as progression decreases asymptotically.

It may be that the observed linearity in the rate of growth is range-bound because of the small number of patient visits, and applies to the slope in a nonlinear growth model. One such possible nonlinear growth model is the power law function, but this was ruled out by the results of the current investigation. The other possible candidate for a nonlinear growth model is a logistic or sigmoidal function, especially in the case of a single lesion, but possibly also for multifocal lesions and total GA (Fig. 6). The data from clinical presentations are typically limited, however, and fail to show either a toe or shoulder region in the scatterplot.

It is plausible that the toe-end of the sigmoidal function could be missing because of the late presentations of patients (i.e., in which vision loss as a result of GA is not apparent at first and patients delay clinical consultations).^{3,30} The shoulder-end of a hypothetical sigmoidal function may also be missing because of the late age and mortality of the patients (i.e., GA does not reach the physical limit of the retina area

or when growth tapers asymptotically). These assumptions are consistent with clinical observations that there is an absence of data at the start and end of the GA data time-series that may indicate the presence of a sigmoidal function. In practice, there are usually only a small number of patient visits and the linear approximation may provide a reliable estimate of trend and a metric for GA progression. The key parameter is the gradient, g , where a high value represents a fast progressor and a low gradient a slow progressor.

Regression models constructed using paired eyes from the same population may have results affected by intereye correlation if treated as a single probability distribution for drawing inferences. Ying et al.³¹ reported methods for addressing this problem for a limited sample based on using left or right eye data only, or random eye selection from each pair, but this reduces sample size. Ying et al.³¹ also suggested the application of mixed-effects models, but with increased computational expense or processing. It was noted that, if there is intereye correlation present and it is ignored, the estimators of the regression coefficients are unbiased, but the variances of the estimators are not correct.³¹ This means that the model average is not affected, but the variance is likely to be smaller than otherwise. The impact of intereye correlation in the current study was expected to be marginal, however, because the regression models developed for each eye in the comparison were derived independently for each eye (not from the total eye population). The regression coefficients were based only on the GA area progression scatter plot for each eye separately. After repeating the process for all eyes, the performance metric was the average value of the r^2 values of the set of models, not the variance. The median of the r^2 values for the models was also computed for the best-fit three models (linear, logarithmic, and quadratic Q2), revealing similar results to the average, with no change to the conclusions (significant because the median involves a ranking of model outputs).

Finally, further subanalysis on left eye data only for paired-eyed patients revealed similar results to Tables 1 and 2. That is, of the best of the three models, none was better than the linear model with respect to average r^2 , and the linear model was still lowest with respect to average U for all models.

Future Extensions

In the current study, there was no evidence found that any of the alternative models tested was better than the linear model for representing total GA growth. The square root transformation showed only marginal differences when compared with untransformed GA

patterns of growth. The models of overall growth for segmented lesions used the total data available, but it is possible that there are subgroups of lesions that may be more appropriately described by different models. For example, recent publications reported that subretinal drusenoid deposits (SDD) may enhance the risk of developing GA, as well as accelerate GA growth. Future investigations could include GA subgroup analysis (such as SDD presence/absence) and the effect on GA growth.^{32,33} Extensions could include further study of the square root transformation applied to GA subgroups using the best three regression models in the comparison (linear, logarithmic, and Q2 models).

Conclusions

Epistemic uncertainty relates to the lack of information on error sources affecting GA measurements that are not associated with statistical variability or replications. An example of an important epistemic uncertainty is “model structure uncertainty” for representing the growth of total GA. A case study on correct model structure was undertaken to compare linear and nonlinear models for representing the trend in GA progression (cf. linear, exponential, logarithmic, quadratic, and power models). Clinical data used in this study consisted of FAF images segmented by the Spectralis HRA+OCT instrument and the Region-Finder algorithm.

The study found no evidence that any other model tested in the comparison was better than using the linear model for describing growth of total GA in the retina. The linear model provides a simple, reliable, and easily interpreted model for charting the trend of GA progression within the range of measurements encountered in clinical practice. The model is compatible with the typically small number of patient visits and time-series data available in clinical presentations. The model was shown to be a balance between statistical significance and clinical assumptions for describing the growth of total GA.

The gradient parameter of the linear model can be interpreted as the rate of GA progression and is an objective metric to evaluate proposed interventions. A low value of the gradient identifies a slow progressor, and a high value identifies a fast progressor. Following an intervention, a new regression model is applied to compute the new gradient. It is suggested that the gradient of the linear model may be the slope of a growth model, such as a sigmoidal function. For different reasons, the toe and shoulder regions are not captured in the time-series data in clinical settings. A

discussion is provided on how the results for the linear approximation reconcile statistical performance with growth assumptions from clinicians.

Acknowledgments

Supported by the National Health & Medical Research Council of Australia (NHMRC) Senior Research Fellowship 1138585 (PNB); NHMRC Fellowship GNT1103013 (RHG); and RB McComas Research Scholarship in Ophthalmology from the University of Melbourne (JA).

Disclosure: **J. Arslan**, None; **K.K. Benke**, None; **G. Samarasinghe**, None; **A. Sowmya**, None; **R.H. Guymner**, is on the advisory boards of Bayer, Novartis, Roche Genentech, and Apellis; **P.N. Baird**, None

References

- Boyer DS, Schmidt-Erfurth U, van Lookeren Campagne M, et al. The pathophysiology of geographic atrophy secondary to age-related macular degeneration and the complement pathway as a therapeutic target. *Retina*. 2017;37(5):819–835.
- Holz FG, Strauss EC, Schmitz-Valckenberg S, van Lookeren Campagne M. Geographic atrophy: clinical features and potential therapeutic approaches. *Ophthalmology*. 2014;121(5):1079–1091.
- Solomon SD, Lindsley K, Vedula SS, et al. Anti-vascular endothelial growth factor for neovascular age-related macular degeneration. *Cochrane Database Syst Rev*. 2014(8):CD005139.
- 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(4):295–317.
- 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.
- Shen LL, Sun M, Khetpal S, et al. 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):2.
- Dreyhaupt J, Mansmann U, Pritsch M, et al. Modelling the natural history of geographic atrophy in patients with age-related macular degeneration. *Ophthalmic Epidemiol*. 2005;12(6):353–362.
- Lindblad AS, Lloyd PC, Clemons TE, et al. Change in area of geographic atrophy in the Age-Related Eye Disease Study: AREDS report number 26. *Arch Ophthalmol*. 2009;127(9):1168–1174.
- Brzozek C, Benke KK, Zeleke BM, et al. Radiofrequency electromagnetic radiation and memory performance: sources of uncertainty in epidemiological cohort studies. *Int J Environ Res Public Health*. 2018;15(4):592.
- Regan HM, Colyvan M, Burgman MA. A taxonomy and treatment of uncertainty for ecology and conservation biology. *Ecol Appl*. 2002;12(2):618–628.
- Limbourg P. Multi-objective optimization of problems with epistemic uncertainty. In: Coello CA, Hernández Aguirre A, Zitzler E, eds. *Evolutionary Multi-Criterion Optimization*. EMO 2005. Lecture Notes in Computer Science, vol. 3410. Berlin, Heidelberg:Springer; 2005.
- Draper D. Assessment and propagation of model uncertainty. *J R Stat Soc Series B Stat Methodol*. 1995;57(1):45–97.
- Elith J, Burgman MA, Regan HM. Mapping epistemic uncertainties and vague concepts in predictions of species distribution. *Ecol Modell*. 2002;157(2):313–329.
- Csaky KG, Richman EA, Ferris FL, 3rd. Report from the NEI/FDA Ophthalmic Clinical Trial Design and Endpoints Symposium. *Invest Ophthalmol Vis Sci*. 2008;49(2):479–489.
- Delori FC, Dorey CK, Staurenghi G, et al. In vivo fluorescence of the ocular fundus exhibits retinal pigment epithelium lipofuscin characteristics. *Invest Ophthalmol Vis Sci*. 1995;36(3):718–729.
- Lindner M, Nadal J, Mauschwitz MM, et al. Combined fundus autofluorescence and near infrared reflectance as prognostic biomarkers for visual acuity in foveal-sparing geographic atrophy. *Invest Ophthalmol Vis Sci*. 2017;58(6):BIO61–BIO67.
- Yung M, Klufas MA, Sarraf D. Clinical applications of fundus autofluorescence in retinal disease. *Int J Retina Vitreous*. 2016;2:12.
- Pollreisz A, Neschi M, Sloan KR, et al. Atlas of human retinal pigment epithelium organelles significant for clinical imaging. *Invest Ophthalmol Vis Sci*. 2020;61:13.
- Brinkmann CK, Schmitz-Valckenberg S, Goebel AP, et al. Semi-automated image analysis software for quantification and calculation of geographic atrophy progression. *Invest Ophthalmol Vis Sci*. 2011;52(14):161.

20. Chaikitmongkol V, Tadarati M, Bressler NM. Recent approaches to evaluating and monitoring geographic atrophy. *Curr Opin Ophthalmol*. 2016;27(3):217–223.
21. Panthier C, Querques G, Puche N, et al. Evaluation of semiautomated measurement of geographic atrophy in age-related macular degeneration by fundus autofluorescence in clinical setting. *Retina*. 2014;34(3):576–582.
22. Pfau M, Goerdt L, Schmitz-Valckenberg S, et al. Green-light autofluorescence versus combined blue-light autofluorescence and near-infrared reflectance imaging in geographic atrophy secondary to age-related macular degeneration. *Invest Ophthalmol Vis Sci*. 2017;58(6):BIO121–BIO130. Erratum in: *Invest Ophthalmol Vis Sci* 2018;59(2):674.
23. Yehoshua Z, et al. Progression of geographic atrophy in age-related macular degeneration imaged with spectral domain optical coherence tomography. *Ophthalmology*. 2011;118:679–686.
24. Pfau M, Lindner M, Goerdt L, et al. Prognostic value of shape-descriptive factors for the progression of geographic atrophy secondary to age-related macular degeneration. *Retina*. 2019;39:1527–1540.
25. Monés J, Biarnés M. The rate of progression of geographic atrophy decreases with increasing baseline lesion size even after the square root transformation. *Transl Vis Sci Technol*. 2018;7(6):40.
26. Domalpally A, Danis RP, White J, et al. Circularity index as a risk factor for progression of geographic atrophy. *Ophthalmology*. 2013;120(12):2666–2671.
27. Julious SA. Sample size of 12 per group rule of thumb for a pilot study. *Pharmaceut Statist*. 2005;4(4):287–291.
28. Johanson GA, Brooks P. Initial scale development: sample size for pilot studies. *Educ Psychol Meas*. 2010;70(3):394–400.
29. Robinson NJ, Benke KK, Norng S. Identification and interpretation of sources of uncertainty in soils change in a global systems-based modelling process. *Soil Res*. 2015;53(6):592–604.
30. Augustin AJ, Kirchhof J. Inflammation and the pathogenesis of age-related macular degeneration. *Expert Opin Ther Targets*. 2009;13(6):641–651.
31. Ying GS, Maguire MG, Glynn R, Rosner B. Tutorial on biostatistics: linear regression analysis of continuous correlated eye data. *Ophthalmic Epidemiol*. 2017;24(2):130–140.
32. Reiter GS, Told R, Schranz M, et al. Subretinal drusenoid deposits and photoreceptor loss detecting global and local progression of geographic atrophy by SD-OCT imaging. *Invest Ophthalmol Vis Sci*. 2020;61:11.
33. Thiele S, Nadal J, Pfau M, et al. Prognostic value of intermediate age-related macular degeneration phenotypes for geographic atrophy progression. *Br J Ophthalmol*. 2021;105:239–245.