tvst

Article

Progression of Geographic Atrophy: Epistemic Uncertainties Affecting Mathematical Models and Machine Learning

Janan Arslan[1,2] and Kurt K. Benke[3,4]

[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

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: June 2, 2021
Accepted: September 27, 2021
Published: November 2, 2021

Keywords: age-related macular degeneration; geographic atrophy; epistemic uncertainty

Citation: Arslan J, Benke KK. Progression of geographic atrophy: Epistemic uncertainties affecting mathematical models and machine learning. Transl Vis Sci Technol. 2021;10(13):3, https://doi.org/10.1167/tvst.10.13.3

Purpose: The purpose of this study was to identify a taxonomy of epistemic uncertainties that affect results for geographic atrophy (GA) assessment and progression.

Methods: An important source of variability is called "epistemic uncertainty," which is due to incomplete system knowledge (i.e. limitations in measurement devices, artifacts, and human subjective evaluation, including annotation errors). In this study, different epistemic uncertainties affecting the analysis of GA were identified and organized into a taxonomy. The uncertainties were discussed and analyzed, and an example was provided in the case of model structure uncertainty by characterizing progression of GA by mathematical modelling and machine learning. It was hypothesized that GA growth follows a logistic (sigmoidal) function. Using case studies, the GA growth data were used to test the sigmoidal hypothesis.

Results: Epistemic uncertainties were identified, including measurement error (imperfect outcomes from measuring tools), subjective judgment (grading affected by grader's vision and experience), model input uncertainties (data corruption or entry errors), and model structure uncertainties (elucidating the right progression pattern). Using GA growth data from case studies, it was demonstrated that GA growth can be represented by a sigmoidal function, where growth eventually approaches an upper limit.

Conclusion: Epistemic uncertainties contribute to errors in study results and are reducible if identified and addressed. By prior identification of epistemic uncertainties, it is possible to (a) quantify uncertainty not accounted for by natural statistical variability, and (b) reduce the presence of these uncertainties in future studies.

Translational Relevance: Lowering epistemic uncertainty will reduce experimental error, improve consistency and reproducibility, and increase confidence in diagnostics.

Introduction

Geographic atrophy (GA) is a debilitating eye disease affecting 5 million individuals globally with expected growth to reach approximately 9 to 10 million individuals by the year 2040.[1] GA appears as lesions which are the result of dead retinal pigment epithelium (RPE) and photoreceptor cells with closure of the underlying choriocapillaris.[2,3] The presence of these lesions in the retina can cause irreversible vision loss, and the size and location of the lesions in the macula is linked with the degree of vision loss.[4,5] The rate of progression of GA is highly variable and there is continuing research on possible factors that contribute to GA and its progression.[5]

There is currently no objective, quantitative, and universally agreed model for progression.[1,5–7] A lack of consensus may be due to the unaccounted variability in many study findings, which is attributable in part to uncertainties associated with the accuracy and precision of various assessment methods.[8] Table 1 summarizes the common epistemic uncertainties that occur in the analysis of GA in research and clinical practice.

Aside from the impact on clinical diagnosis and management, uncertainty analysis is important

Table 1. Epistemic Uncertainties in the Analysis of GA

Sources of uncertainty
• Data quality (data entry errors, artifacts, noise, duplication, and data corruption)
• Image quality (contrast, resolution, color, optical aberrations, and sensor noise)
• Measurement error (instrumentation resolution and reproducibility)
• Grader – annotation errors and population variability
• Education and experience
• Visual acuity and concentration
• Fatigue and stress
• Model structure uncertainty (choice of model and parameters)
• Model parameter uncertainty (regression analysis and confidence intervals)
• Sample size and homogeneity (statistical significance)
• Choice of biophysical model versus machine learning model

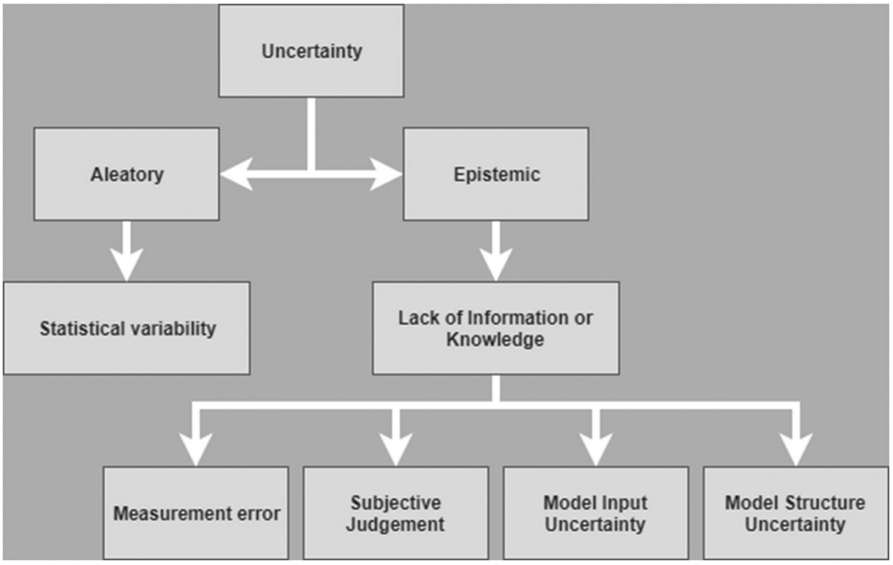

Figure 1. A taxonomy of uncertainty highlighting the different types of epistemic uncertainty that occur in GA assessment in addition to statistical variability associated with replications. GA, geographic atrophy.

because progression models and machine learning can be affected by data quality and human annotation errors during the course of training and parameter estimation. Some GA analytic models are hybrid approaches combining features of biophysical approaches and machine learning. These include logistic models and mixed-effects models.

Identification of epistemic uncertainties could (a) statistically quantify variability not accounted for by a regression model, and (b) provide information for reducing these uncertainties (e.g. by experimental modification, data normalization, and image preprocessing).

In the taxonomy of uncertainty, there are two broad categories of classification: aleatory uncertainty and epistemic uncertainty (Fig. 1). Aleatory uncertainty is regarded as irreducible uncertainty and is the natural statistical variation in data and experimental studies.[9] Epistemic uncertainty is due to lack of knowledge and refers to reducible errors, such as subjective uncertainty, measurement error, and model structure uncertainty. Epistemic uncertainty can also arise due to the limitations of electronic instrumentation and corrupted data.[10,11] By identifying significant epistemic uncertainties, statistical techniques can be used to reduce their impacts on the assessment of GA.

In a previous publication by the authors, various GA progression models were evaluated in a study of model structure uncertainty.[12] Other types of epistemic

uncertainty were not investigated. Subsequently, an online search revealed that epistemic uncertainty in GA assessment in age-related macular degeneration (AMD) appears to be a neglected area of research. No other publications were found on epistemic uncertainty in GA assessment using fundus autofluorescence images apart from the prior work by the authors. In the current study, we performed a taxonomic analysis to identify and categorize other sources of epistemic uncertainty. In addition, one hypothesis from the previous paper was also tested (i.e. that although the linear approximation is generally apparent and sufficient in most clinical applications, the entire process of GA progression from start to completion may actually follow a sigmoidal model).[12] The hypothesis was investigated as a subanalysis of the data in the previous paper, for subjects with a sufficient number of clinical presentations.

Methods

Taxonomy of Uncertainty

Epistemic uncertainties are sources of imprecision that affect GA assessment and hamper the development of new and suitable model designs for GA growth.[11,13,14] The presence of these uncertainties, if sufficient in magnitude, can affect the quality of the information collected from experiments, leading to variability in results, lack of reproducibility, and lower prediction accuracy. Epistemic uncertainties relevant to GA progression were identified in the assessment process, such as measurement error, subjective judgment, model input uncertainty, and model structure uncertainty. The sources of epistemic uncertainty were organized into a taxonomy and expressed as a process flowchart to assist in identification and possible intervention by modifications in experimental designs and data collection.

The epistemic uncertainties described in this investigation are relevant to fundus autofluorescence (FAF) images acquired by the Spectralis HRA + OCT instrumentation (Heidelberg Engineering) and the associated RegionFinder segmentation software. The study provides insights and information relevant to other imaging modalities in ophthalmology, such as color fundus photographs.

Case Study: Model Structure Uncertainty

The patients in this subanalysis for model structure uncertainty were selected from a GA-affected cohort collected retrospectively and used in a recent prior study on GA assessment and progression.[12] This study was approved by the Human Research Ethics Committee of the Royal Victorian Eye and Ear Hospital (RVEEH) and 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. The cohort was previously used in the evaluation of GA progression using quantified epistemic uncertainty for model structure analysis and consisted of 81 eyes from 45 patients.[12] The number of images per eye were acquired from patients for 3 to 17 visits (i.e. clinical consultations).

For the subanalysis, patients with a high number of clinical presentations were selected for further investigation. The choice was pragmatic, based on finding patients with the most visits to provide sufficient data for regression analysis to test a proposed sigmoidal growth model. Four patients were found in the cohort with a double-digit number of clinical presentations suitable for modeling (ranging between 11 and 17 visits). Further information on the data origins for the study can be found in the publication by Arslan et al. (2021).[12]

Results

Taxonomy of Uncertainty in GA Assessment

In Figure 2, a flowchart shows where epistemic uncertainties can occur in modeling the GA growth process. In this section, more details are provided on the issues of image processing, data quality, and model structure uncertainty, so that more information is available to inform future research in GA assessment and progression.

Measurement Error

Measurement errors are associated with limitations and imperfections in the instrumentation, including sensor resolution, reproducibility, electronic noise, artifacts, and distortion. A warm-up time may be needed for laboratory instrumentation after a cold start, and there may be batch-to-batch differences in equipment and differences between manufacturers due to optics or electronics. All of these errors are potentially reducible.

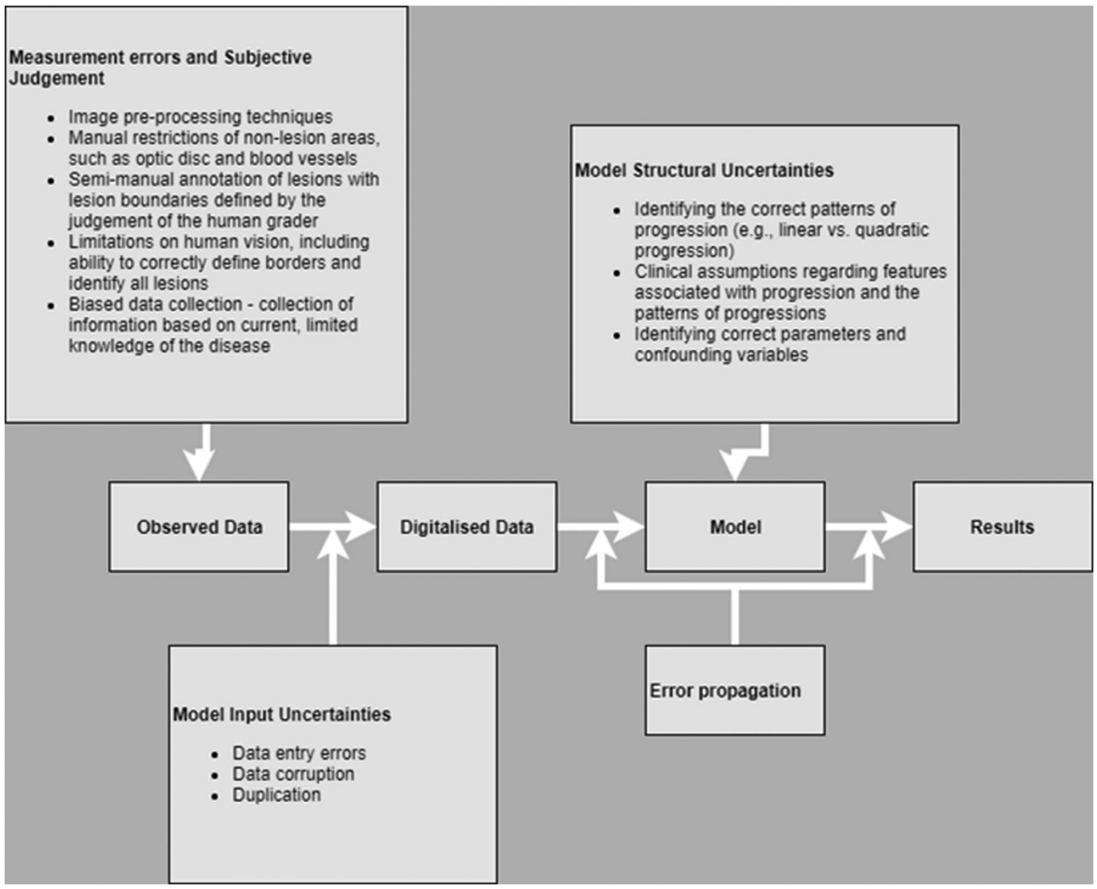

Figure 2. Principal uncertainties in the modelling process - from data to results.

Bias in Data

Data collection is based on past knowledge of the disease and thus there is epistemic uncertainty in the relevant predictors, such as patient lifestyle and medical history. There may be a lack of quality assurance in the data type and format and whether it is appropriate for investigation of the disease. The number of covariates investigated may also be limited by the limited availability of patient data. For example, the association between smoking and GA progression has been a long-standing discussion, but not all institutions have available smoking data or common formats.[15] There may be drivers of GA and its progression not previously recognized and therefore past data collected is flawed or insufficient.

Model Input Uncertainty

Input errors for predictive models can be due to data entry errors (e.g. incorrect entry of dates for patient visits), transferring software data into spreadsheets (e.g. exporting RegionFinder results into another database), data corruption (e.g. issues in reading, writing, and storing data), and duplication (e.g. multiple entries pertaining to the same data point). Data quality can be checked and uncertainty is reducible with rigorous quality assurance and data cleansing procedures that systematically check for duplications, negative numbers, or impossible dates.

Subjective Uncertainty

Subjective judgment can also lead to uncertainties due to bias in expert opinion.[16–18] Subjective uncertainties include human judgment used to manually restrict non-lesion areas, limitations in human vision in correctly identifying lesions and their boundaries, limitations in the current understanding of GA atrophy, assumptions surrounding the progression of the disease, and bias in the data collection process. Repeatability (of the same experiment) and reproducibility (by others) of human-defined annotations are associated with intraobserver and interobserver variability (Figs. 3a, 3b).[19] Figure 4 is an example of an FAF image annotated at different time points

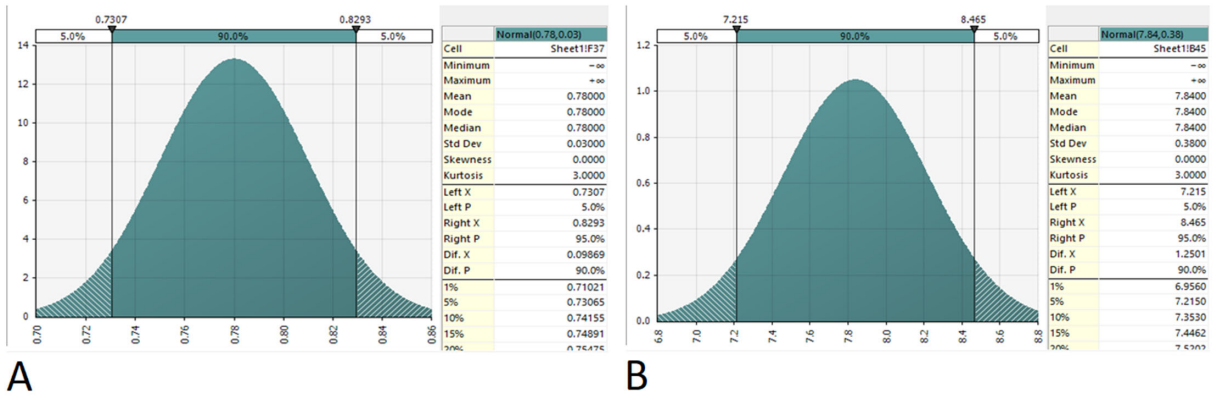

Figure 3. (**A**) Example of interoperator variability with respect to Dice coefficients. Plotted curve produced using the technique of statistical bootstrapping (Benke et al. 2018) applied to data from Liefers et al. (2020).[27,28] (**B**) Shows an example of possible intraoperator variability with respect to segmentation in fundus autofluorescence images. Plotted curve produced using the technique of statistical bootstrapping applied to experimental data (Benke et al. 2018).[27]

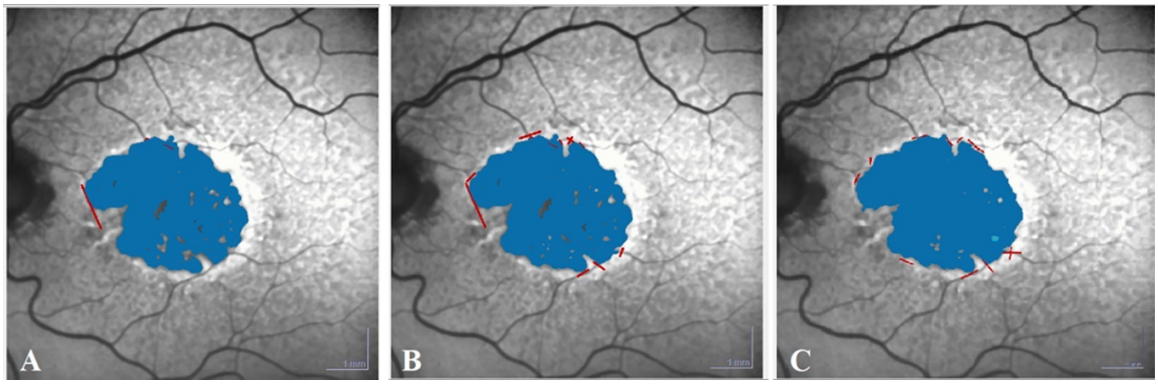

Figure 4. Example of potential intraobserver variability. These annotations are for the same image conducted by the same grader. The total area for the original annotation (**A**) was 7.497 mm^2. The total areas for the second (**B**) and third (**C**) annotations were 7.770 mm^2 and 8.248 mm^2, respectively.

by the same grader. Annotation errors in GA by the grader can be due to limitations in human vision, knowledge, experience, expectations, fatigue, and stress.

Optics and Image Processing

Errors in the FAF image acquisition process include the appearance of dark contrasts, nonuniform illumination, misaligned image orientation, and blurriness caused by microsaccades, which include small and subtle involuntary eye movements. Poor image quality can further contribute to the development of uncertainty, as artifacts may interfere with the observations of a grader, who may miss important features in the image structure.[20]

RegionFinder software can be used to retrospectively assess GA progression and has preprocessing operations to address image artifacts, image registration, shadow correction, and speckle noise removal.[21] The use of image registration (i.e. alignment of image features, such as the optic disc and blood vessels, to the same coordinates for future image comparisons) and shadow correction ensure accuracy of capturing disease changepoint detection. However, speckle-noise removal, while reducing granularity in FAF imagery can inadvertently highlight existing features, such as expanding the size of the blood vessels. It can also remove important information on granularity-like features of the disease that are not image-acquired noise but physiological markers of the disease. In place of standard speckle-noise removal methods, filters specifically designed to remove granularity noise while preserving information pertaining to GA would be more suitable, such as the median filter.[22] Preserving retinal granularities could be useful in understanding the appearance and progression of lesions and even hyperfluorescent areas.

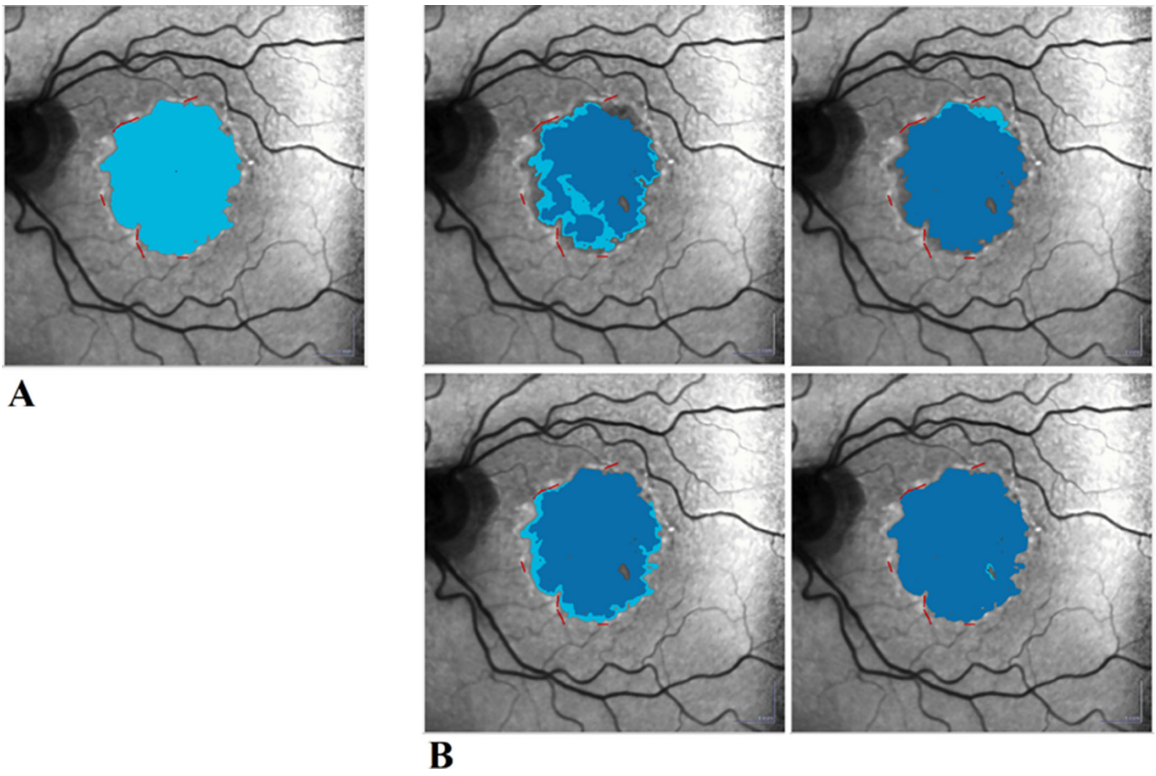

Figure 5. Singular seed versus multiple seed GA annotation. Although a user could annotate a single lesion from a single seed point, as illustrated in (**A**), in some instances, this is not feasible when the shapes are not entirely circular. Instead, users can opt to use multiple seeds in their GA annotations, as shown in (**B**). This option also allows for greater control in annotating with greater accuracy (e.g. correctly annotating boundaries). GA, geographic atrophy.

Expert Opinion - Identification of Lesion Boundaries and Perimeters

The RegionFinder software utilizes a region-growing algorithm for lesion segmentation and annotation (see Fig. 5). It requires the user to select a "seed" point within the lesion area, and then, using the mouse cursor, the seed can be expanded to encompass the entire area of the lesion based on color similarity. Graders have noted that using a single seed point is not always sufficient and multiple seed points may be needed to effectively cover the area (Fig. 5 illustrates a lesion annotation using a single seed point versus multiple seed points). This issue is coupled with the ability of the end-user to define lesion borders correctly and stop expanding the seed. Smaller and more discrete lesions may be missed during the annotation process.

Expert Opinion – Manual Restrictions of Non-Lesion Areas

As the optic disc, blood vessels, and fovea all have similar color intensities to that of GA lesions, the region-growing algorithm could "spill" from the lesion into ocular features with similar intensity (Fig. 6). The RegionFinder software has a restriction function that allows the end-user to limit expansion of seeds into non-lesion areas. Additionally, RegionFinder allows the user to replicate these restrictions in future images to save time and ensure consistency. This feature requires human-user input to correctly differentiate between ocular features, introducing uncertainty.

Model Structure Uncertainty

An important epistemic uncertainty associated with modeling time-series data, such as the progression of GA atrophy, is "model structure uncertainty." This error source is distinct from "parameter uncertainty," which is associated with model calibration by regression analysis.[12,23] The structure of the model is an important source of error as it is likely to have implications in both prediction accuracy and confidence intervals produced by the model.

Some methods available to address structural problems include model checking (e.g. goodness-of-fit

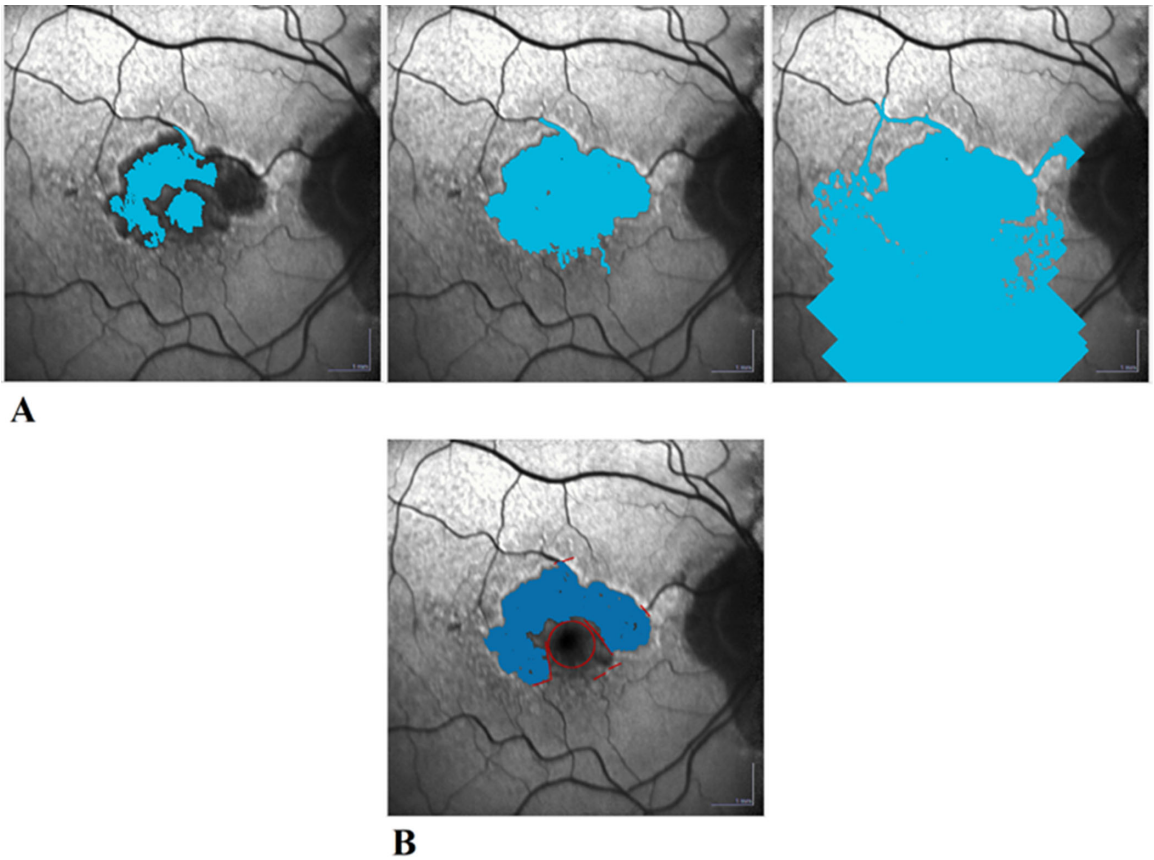

Figure 6. Manual restriction of non-lesion areas. (**A**) Step-by-step lesion annotation without any manual restrictions on the fovea and blood vessels, and (**B**) complete lesion annotation with manual restrictions on fovea and blood vessels. Manual restrictions exist to ensure that the region-growing algorithm does not "spill" into non-lesion areas.

tests, calibration test, and residual error assessment) and comparing tested predictions against independent data.[17]

In the calibration of a regression model, unexplained variability due to uncertainty can be expressed as a function of the coefficient of determination, r^2:

$$U = 1 - r^2,$$

where

$$r^2 = 1 - \frac{SSR}{SSO}$$

where the ratio $\frac{SSR}{SSO}$ = sum of square residuals divided by the total sum of squares in the data.

The metric U is the proportion of total unexplained variability not accounted for by the regression model. Traditional statistical regression analysis assumes that the residuals arise from statistical variability, while ignoring the contribution from epistemic uncertainty. Ideally, a model with the lowest U and highest r^2 would be considered the most suitable for a GA growth model.

In a recent paper by Arslan et al. (2021) on uncertainty in characterization and growth of GA, it was found in a statistical comparison of regression models (cf. power law function, logarithmic, exponential, and quadratic models), that no model tested performed any better than the linear model for characterizing GA growth.[12,23] It was hypothesized that the linear model was the slope of a growth model, such as a logistic (sigmoidal) function, but that there was insufficient data from typical clinical presentations to elicit this functional form. The linear model provided an objective metric for the rate of GA progression in the form of the gradient, which could be used to compare interventions. For the current study, a search of our electronic health record (EHR) database found patients with sufficient data to investigate further the issue of model structure uncertainty. The analyses are described in the following retrospective case studies using anonymized data.

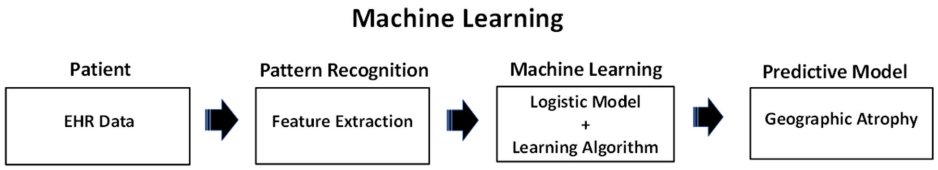

Figure 7. Machine learning in the case study was characterized by a process of pattern recognition and parameter estimation which produced a model for GA progression. GA, geographic atrophy.

Case Study – Modeling Progression of GA

Patient X in the subanalysis was an elderly female and the first case investigated. The GA area measurements were acquired over 8 years (i.e. June 2010 to April 2018) in a total of 16 visits and stored as EHR data. The second case (Patient Y) was an elderly female patient at the time of initial presentation. The patient attended from November 2011 to February 2019 for a total of 17 visits. The third case study (Patient Z) was also an elderly female patient at initial presentation. The patient was seen from June 2012 to June 2019 for a total of 11 visits. The fourth case study (Patient W) was a middle-aged male patient at initial presentation and attended between February 2010 and December 2017 for 14 visits in total. The unusually large number of clinical presentations for the four patients provided sufficient data for testing a multiparameter nonlinear model for progression of GA.

Although a linear model is effective for small data sets and provides a gradient as a metric for rate of growth of GA in a clinical setting, a nonlinear model may provide additional information on disease onset and end point, if larger datasets are available.[24] A growth curve with turning points can be represented by a logistic function, such as the sigmoidal model,[24]

$$G = C + \left[\frac{A}{1 + \exp(-B(t - M))} \right]$$

where G is geographic atrophy at time t (i.e. date, or number of clinical presentation). This function is a regression model with the characteristic parameters of a sigmoidal model, such as slope (B), inflection point (M), lower bound (C), and upper bound (A). It is analogous to the classical hill-slope models used for dose-response curves.[24] Notably, it is a growth function characteristic of many biophysical models used in the life sciences.

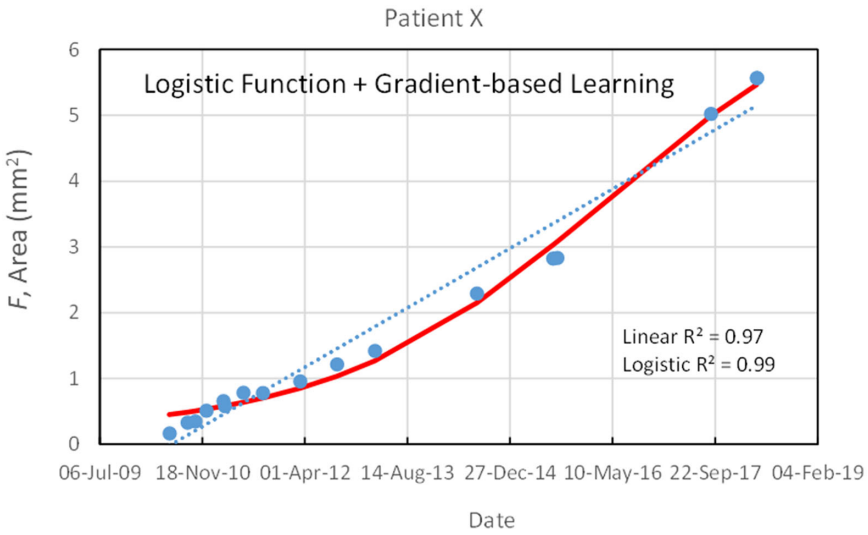

Figure 8. Example of a logistic regression model for GA progression with parameter estimation by a multi-start gradient-based learning approach showing close fit to patient data, $R^2 = 0.99$ ($P < 0.01$), see Table 2. GA, geographic atrophy.

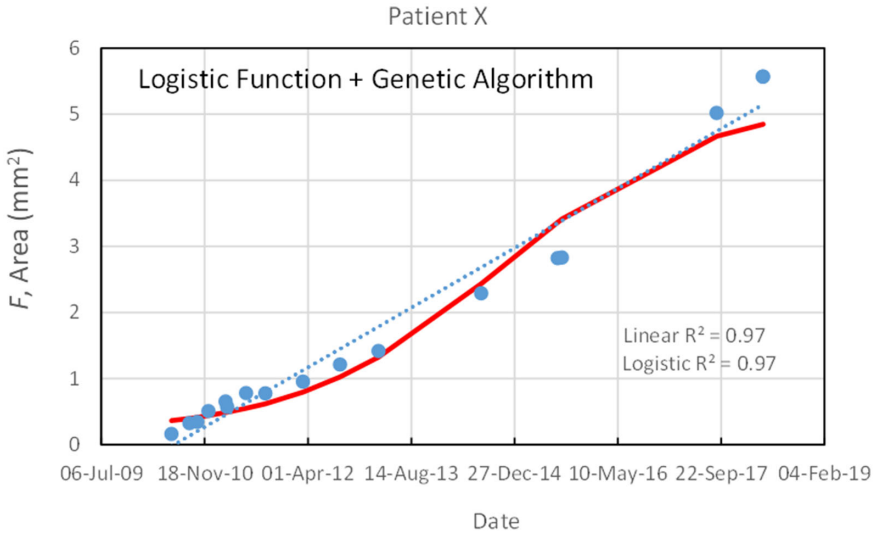

Figure 9. Example of a logistic regression model for GA progression with parameter estimation by a genetic algorithm showing a close fit to the patient data - but rolling-off near the shoulder of the curve. In this case, the fit is less effective than the gradient-based learning model in Figure 8. Again, the gradient of the linear approximation is similar to the gradient at the inflection point of the logistic function. GA, geographic atrophy.

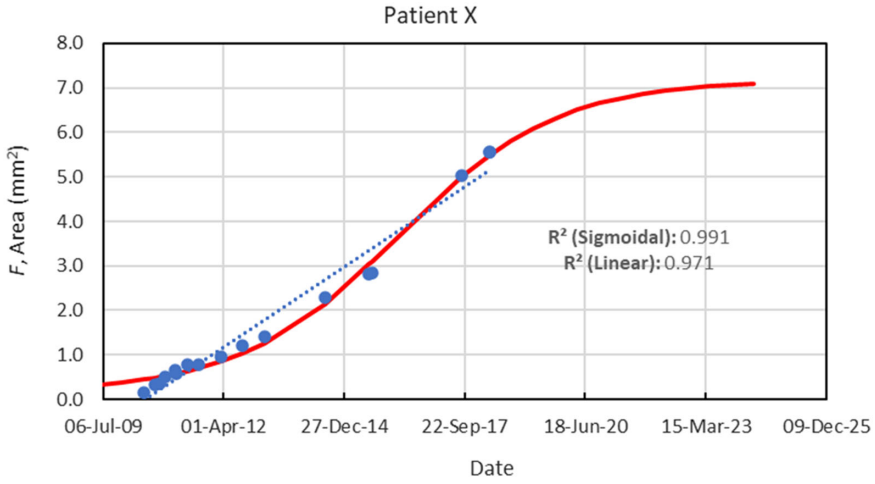

Figure 10. Using the fitted model in Figure 8, extrapolation of the trend in GA reveals it tapers off at about 7 mm^2 after 2023. At the toe of the curve, it appears that GA begins at about Jan 2009 (where it is less than about 1% of peak GA area), providing potentially useful information for research on nascent GA. GA, geographic atrophy.

Figure 7 shows a flowchart for a machine learning approach for a GA progression model characterized by parameter estimation using patient data. The weights of the model would be adjusted iteratively by a proposed learning algorithm until the performance index is optimized.[25] A nonlinear regression analysis can be used to find the optimum parameter set.

Machine learning approaches tend to lead to more complex models than biophysical models.

Figure 8 for Patient X shows a logistic function model for GA progression with parameter estimation carried out by a gradient-based learning approach,[24] based on initialization at 10 random starting points. Results for nonlinear least-squares regression analysis suggested convergence to a probable global solution on the error surface of residuals ($R^2 = 0.99$, $P = 0.01$ for r). Note that the gradient of the linear approximation (broken line) is similar to the gradient at the inflection point of the logistic function (see also Figure 5 in Choi et al. [2020]).[25] The gradient has been proposed as a convenient metric for rate of progression in clinical applications.[12]

Figure 9 for Patient X shows a logistic function model for GA progression with parameter estimation by a genetic optimization algorithm. This type of algorithm is used in computer science for global

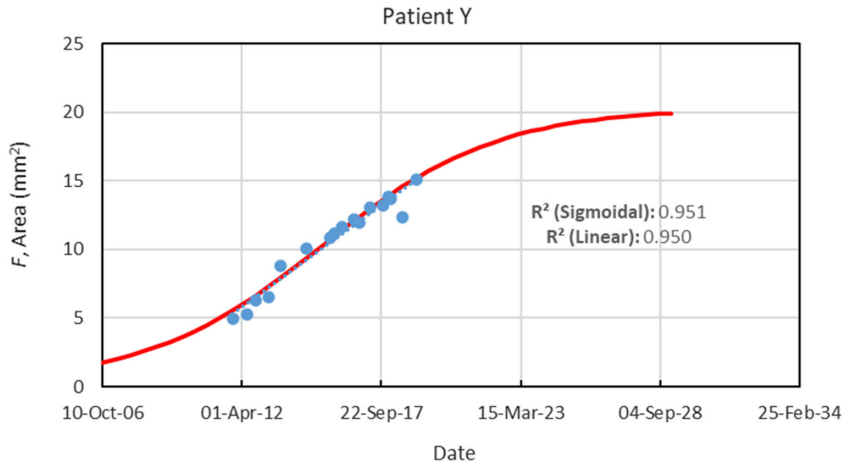

Figure 11. For Patient Y, comparison between linear and sigmoidal models shows a close match over mid-range ($R^2 = 0.95$ for both models). Extrapolation of the sigmoidal fit suggests GA tapers off at about 20 mm^2 after 2028. At the toe of the curve, it appears GA begins at about June 2006. GA, geographic atrophy.

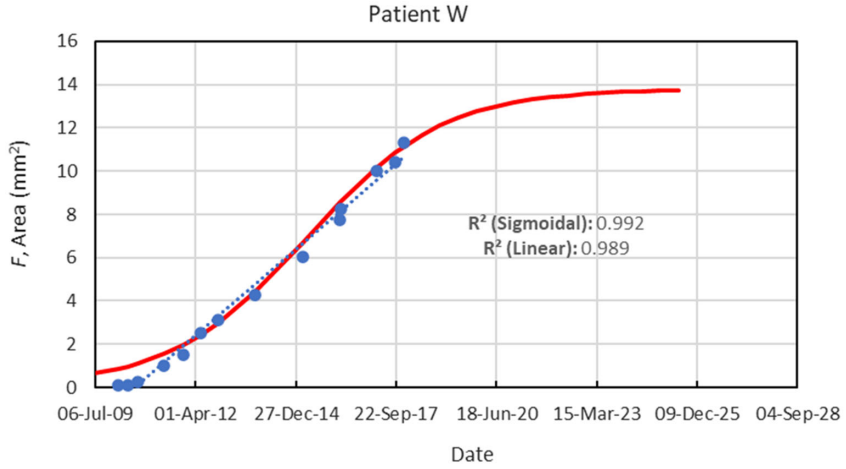

Figure 12. For Patient W, comparison between linear and sigmoidal models shows a close match over mid-range ($R^2 = 0.99$ for both models). Trend in GA from the sigmoidal model reveals tapering off at about 14 mm^2 after 2023. At the toe of the curve, it appears that GA begins at around June 2006. GA, geographic atrophy.

optimization and is based on reinforcement rules inspired by an evolutionary programming approach (population size 10, mutation rate 0.075, and max time without change, 30 trials). The results show a close fit to the patient data - but rolling-off near the shoulder of the curve. The fit is less effective in this case than the gradient-based learning model. Again, the gradient of the linear approximation is similar to the gradient at the inflection point of the logistic function.

Figure 10 for Patient X shows extrapolation of the fitted model (previously plotted in Fig. 8), revealing that the trend in GA for Patient X tapers off at about 7 mm^2 after 2023. At the toe of the curve, it appears that GA originated at around January 2009 (where it is less than about 1% of the peak GA area). Before this time, there may still have been GA present, but not yet visible in clinical settings. This may provide useful information for research on nascent GA. Note that the broken line shows the slope of a linear regression model for comparison.[24]

Figures 11 to 13 show results for Patients Y, Z, and W, which are consistent with the results for Patient X, showing that the linear model is a good approximation to GA progression for each patient (Table 2). Given sufficient data, however, the sigmoidal model may provide even closer fits to the data (R^2), together with additional information on possible GA onset and tapering to limiting values. The results for the four patients are shown in Table 2.

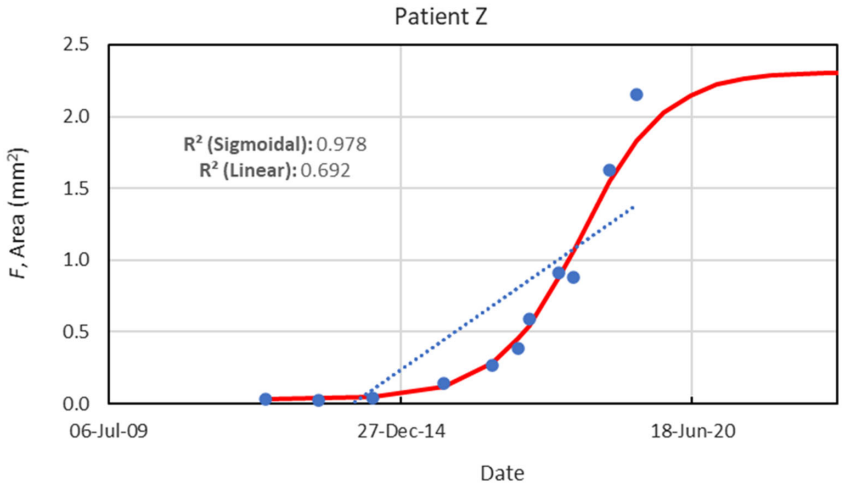

Figure 13. For patient Z, comparison shown between linear and sigmoidal models ($R^2 = 0.69$ for linear model, $R^2 = 0.98$ for sigmoidal model). The strong early curvature evident in the scatter plot means that the linear approximation should be applied more effectively after 2015 (where linear $R^2 = 0.89$). GA, geographic atrophy.

Table 2. Results for Modeling GA Progression

Patient Label	Visits n	Parameters (Sigmoidal)			Metrics			
		C	A	B	R^2 (Sigmoidal)	R^2 (Linear)	U (Sigmoidal)	U (Linear)
X	16	0.17694	7	0.001516	0.991	0.971	0.009	0.029
Y	17	0.27	20	0.0008	0.951	0.950	0.049	0.050
Z	11	0.03242	2.277065	0.00345	0.978	0.692	0.022	0.308
W	14	0.086306	13.7	0.001485	0.992	0.989	0.008	0.011

Notes: $P < 0.001$ for R values; patient Z, scatter plot data after 2015 produced linear $R^2 = 0.8909$.

The insights gained from this retrospective case study relate to uncertainties in the three possible models (linear approximation, genetic algorithm, and logistic function). First, it supports the hypothesis from a recent study that (a) the slope of the linear approximation is similar to the maximum slope of a sigmoidal model,[24] (b) the gradient is a potential metric for rate of progression of GA, and (c) the logistic function may provide additional information on possible onset and endpoint. In practice, many clinical settings are characterized by data paucity for parameter estimation and so favor a linear approximation by regression analysis as an indicator of the rate of GA progression.

Discussion

Epistemic uncertainties in GA assessment can propagate as error sources in the process of data acquisition, diagnosis, and model development. This results in greater variability and wider confidence intervals and therefore less confidence in testing the original experimental hypothesis. Primary sources of epistemic uncertainties are data quality, digital image processing, and data annotation errors. Other sources of epistemic uncertainty include intergrader and intragrader variability (which may be reducible by increased automation), and "model structure uncertainty" when forecasting progression of GA (reducible by selecting the correct progression model). The impact of uncertainties in data quality and annotation accuracy will affect diagnostics by human graders as well as mathematical models for progression and machine learning.

Identification and systematic treatment of specific epistemic uncertainties will assist in reducing experimental variability.[26] For example, with FAF images, improvements are possible by (1) extending speckle-noise removal by the RegionFinder segmentation software and by investigating additional filters, such as the median filter, which may be more selective in discrimination between system noise and natural granularity, (2) applying machine learning to automate lesion segmentation (reducing human subjectivity in

the annotation process), and (3) increasing sample size and the number of feature measurements. These suggested enhancements could improve delineation of GA boundaries and therefore segmentation performance (i.e. improve the resolution of lesion boundaries for improved feature extraction by the human grader or by machine learning). Further reduction in epistemic uncertainty may be possible by using machine learning approaches to find new features for discrimination in the image that may not be readily apparent to human graders. This could result in greater utilization of available data and may even lead to information discovery and insights that were not previously considered.

Further research on model structure uncertainty in GA progression could progressively minimize this source of epistemic uncertainty, whereas the quest for improving model structure may also help to inform and provide clues to the nature of GA growth. The results in this study suggest that for sparse datasets from clinical presentations, a linear approximation appears reasonable for modeling GA progression whereas, given sufficient data, a sigmoidal model may also provide information with respect to GA onset and asymptotic convergence to a plateau.

In summary, epistemic uncertainties affect experimental data quality, image processing results, and data annotations. This extraneous effect can degrade the performance of human graders, mathematical models, and machine learning performance. There are many sources of uncertainty that can be reduced, especially with FAF images, if guided by the taxonomy and analysis presented in this study.

Conclusions

In this study, a number of sources of epistemic uncertainty have been identified in GA assessment and its progression in fundus autofluorescence images. Unlike natural statistical variability associated with experimental error or replications, epistemic uncertainties are reducible because they relate to lack of information. In particular, epistemic uncertainties can be addressed by appropriate experimental design modifications and data quality assurance.

Epistemic uncertainties can affect grader performance and can also affect mathematical models and machine learning approaches because they are both dependent on experimental data for parameter estimation. A limited retrospective case study was included on the issue of "model structure uncertainty" in GA progression and extends the results and conclusions reported in a recent study.[12]

The results for the sigmoidal model are very encouraging and suggest further study as it has the advantage of providing additional information on possible onset of GA and asymptotic progression to a limiting value by extrapolation beyond the time-series data (subject to a specified level of precision). In most clinical applications, there is a limited number of patient presentations, suggesting recourse to a linear approximation for estimating the rate of progression.

In the future, the study of epistemic uncertainty is likely to be a subject of increasing interest to biostatisticians and clinicians because it relates to additional factors that are often neglected while reporting study results subject to natural statistical variability and experimental errors.

Acknowledgments

Disclosure: **J. Arslan,** None; **K.K. Benke,** None

References

1. Boyer DS, Schmidt-Erfurth U, van Lookeren Campagne M, Henry EC, Brittain C. The Pathophysiology of Geographic Atrophy Secondary to Age-Related Macular Degeneration and the Complement Pathway as a Therapeutic Target. *Retina*. 2017;37:819–835.
2. Solomon SD, Lindsley K, Vedula SS, Krzystolik MG, Hawkins BS. Anti-vascular endothelial growth factor for neovascular age-related macular degeneration. *Cochrane Database Syst Rev*. 2014;8:CD005139.
3. Holz FG, Strauss EC, Schmitz-Valckenberg S, van Lookeren Campagne M. Geographic atrophy: clinical features and potential therapeutic approaches. *Ophthalmology*. 2014;121:1079–1091.
4. 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.
5. Fleckenstein M, Mitchell P, Freund KB, et al. The Progression of Geographic Atrophy Secondary to Age-Related Macular Degeneration. *Ophthalmology*. 2018;125:369–390.
6. Dreyhaupt J, Mansmann U, Pritsch M, Dolar-Szczasny J, Bindewald A, Holz FG. Modelling the natural history of geographic atrophy in

patients with age-related macular degeneration. *Ophthalmic Epidemiol*. 2005;12:353–362.
7. 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;127:1168–1174.
8. Shen LL, Sun M, Khetpal S, Grossetta Nardini HK, Del Priore LV. 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;61:2.
9. Brzozek C, Benke KK, Zeleke BM, Abramson MJ, Benke G. Radiofrequency Electromagnetic Radiation and Memory Performance: Sources of Uncertainty in Epidemiological Cohort Studies. *Int J Environ Res Public Health*. 2018;15:592.
10. Regan HM, Colyvan M, Burgman MA. A taxonomy and treatment of uncertainty for ecology and conservation biology. *Ecological Applications*. 2002;12:618–628.
11. Hester P. Epistemic Uncertainty Analysis: An Approach Using Expert Judgment and Evidential Credibility. *Int J Quality, Statistics, and Reliability*. 2012;2012:8.
12. 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;10:2.
13. Hofer E, Kloos M, Krzykacz-Hausmann B, Peschke J, Woltereck M. An approximate epistemic uncertainty analysis approach in the presence of epistemic and aleatory uncertainties. *Reliability Engineering & System Safety*. 2002;77:229–238.
14. Burnham KP, Anderson DR. Model-Selection Uncertainty with Examples. *Model Selection and Inference: A Practical Information-Theoretic Approach.* New York, NY: Springer New York; 1998:118–158.
15. Seddon J, Willett WC, Speizer FE, Hankinson SE. A Prospective Study of Cigarette Smoking and Age-Related Macular Degeneration in Women. *JAMA*. 1996;276:1141–1146.
16. Kiureghian AD, Ditlevsen O. Aleatory or epistemic? Does it matter? *Structural Safety*. 2009;31:105–112.
17. Elith J, Burgman MA, Regan HM. Mapping epistemic uncertainties and vague concepts in predictions of species distribution. *Ecological Modelling*. 2002;157:313–329.
18. Kujala H, Burgman MA, Moilanen A. Treatment of uncertainty in conservation under climate change. *Conservation Letters*. 2013;6:73–85.
19. Dao TT, Tho MC. Uncertainty modeling of input data for a biomechanical system of systems. *2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. Annu Int Conf IEEE Eng Med Biol Soc. 2013;2013:4581–4584.
20. Holz FG, Spaide RF, Bird AC, Schmitz-Valckenberg S. *Atlas of fundus autofluorescence imaging*. Berlin, New York: Springer; 2007.
21. Brinkmann CK, Schmitz-Valckenberg S, Goebel AP, Fleckenstein M, Stratmann N, Holz FG. Semi-automated Image Analysis Software For Quantification And Calculation Of Geographic Atrophy Progression. *Invest Ophthalmol Vis Sci*. 2011;52:161.
22. Broesch JD. Chapter 7 - Applications of DSP. In: Broesch JD (ed), *Digital Signal Processing.* Burlington: Newnes; 2009:125–134.
23. Arslan J, Benke KK, Guymer RH, Baird PN. Understanding the uncertainties in modelling the growth of geographic atrophy in age-related macular degeneration. *Invest Ophthalmol Vis Sci*. 2020;61:2986–2986.
24. Lasdon LS, Waren AD, Jain A, Ratner M. Design and Testing of a Generalized Reduced Gradient Code for Nonlinear Programming. *ACM Trans Math Softw*. 1978;4:34–50.
25. Choi RY, Coyner AS, Kalpathy-Cramer J, Chiang MF, Campbell JP. Introduction to Machine Learning, Neural Networks, and Deep Learning. *Transl Vis Sci Technol*. 2020;9:14.
26. Arslan J, Samarasinghe G, Benke KK, et al. Artificial Intelligence Algorithms for Analysis of Geographic Atrophy: A Review and Evaluation. *Transl Vis Sci Technol*. 2020;9:57.
27. Benke KK, Norng S, Robinson NJ, Benke LR, Peterson TJ. Error propagation in computer models: analytic approaches, advantages, disadvantages and constraints. *Stochastic Environ Res and Risk Assess*. 2018;32:2971–2985.
28. Liefers B, Colijn JM, González-Gonzalo C, et al. A Deep Learning Model for Segmentation of Geographic Atrophy to Study Its Long-Term Natural History. *Ophthalmology*. 2020;127:1086–1096.