

Estimating Ganglion Cell Complex Rates of Change With Bayesian Hierarchical Models

Vahid Mohammadzadeh¹, Erica Su², Sepideh Heydar Zadeh¹, Simon K. Law¹, Anne L. Coleman¹, Joseph Caprioli¹, Robert E. Weiss², and Kouros Nouri-Mahdavi¹

¹ Glaucoma Division, Stein Eye Institute, David Geffen School of Medicine, University of California Los Angeles, Los Angeles, CA, USA

² Department of Biostatistics, Fielding School of Public Health, University of California Los Angeles, Los Angeles, CA, USA

Correspondence: Kouros Nouri-Mahdavi, Glaucoma Division, Stein Eye Institute, David Geffen School of Medicine, University of California Los Angeles, 100 Stein Plaza, Los Angeles, CA 90095, USA. e-mail: nouri-mahdavi@jsei.ucla.edu

Received: October 19, 2020

Accepted: February 21, 2021

Published: April 15, 2021

Keywords: optical coherence tomography; OCT; macula; progression; ganglion cell complex; GCC; longitudinal; bayesian; superpixels

Citation: Mohammadzadeh V, Su E, Heydar Zadeh S, Law SK, Coleman AL, Caprioli J, Weiss RE, Nouri-Mahdavi K. Estimating ganglion cell complex rates of change with Bayesian hierarchical models. *Trans Vis Sci Tech.* 2021;10(4):15. <https://doi.org/10.1167/tvst.10.4.15>

Purpose: Develop a hierarchical longitudinal regression model for estimating local rates of change of macular ganglion cell complex (GCC) measurements with optical coherence tomography (OCT).

Methods: We enrolled 112 eyes with four or more macular OCT images and ≥ 2 years of follow-up. GCC thickness measurements within central 6×6 superpixels were extracted from macular volume scans. We fit data from each superpixel separately with several hierarchical Bayesian random-effects models. Models were compared with the Watanabe–Akaike information criterion. For our preferred model, we estimated population and individual slopes and intercepts (baseline thickness) and their correlation.

Results: Mean (SD) follow-up time and median (interquartile range) baseline 24-2 visual field mean deviation were 3.6 (0.4) years and -6.8 (-12.2 to -4.3) dB, respectively. The random intercepts and slopes model with random residual variance was the preferred model. While more individual and population negative slopes were observed in the paracentral and papillomacular superpixels, superpixels in the superotemporal and inferior regions displayed the highest correlation between baseline thickness and rates of change ($r = -0.43$ to -0.50 for the top five correlations).

Conclusions: A Bayesian linear hierarchical model with random intercepts/slopes and random variances is an optimal initial model for estimating GCC slopes at population and individual levels. This novel model is an efficient method for estimating macular rates of change and probability of glaucoma progression locally.

Translational Relevance: The proposed Bayesian hierarchical model can be applied to various macular outcomes from different OCT devices and to superpixels of variable sizes to estimate local rates of change and progression probability.

Introduction

Glaucoma is a progressive optic neuropathy and is one of the leading causes of blindness worldwide.^{1,2} Detection of glaucoma progression is essential for preventing irreversible visual loss. Earlier detection of glaucoma worsening would allow clinicians to titrate treatment in a timely fashion. Rates of progression vary among patients. Identifying eyes that are progressing rapidly is important for the appropriate allocation of available resources. In some glaucoma eyes, progres-

sive structural changes can be detected earlier than functional changes.^{3–8}

Macular optical coherence tomography (OCT) is an essential tool for the evaluation of central retinal ganglion cells (RGC).^{9–14} There is now evidence supporting the utility of macular OCT imaging for detection of glaucoma progression, especially in later stages of glaucoma.^{2,10,15–17} Most prior studies of macular thickness changes over time have used global outcome measures or thickness measurements in hemiretinal regions. These approaches may fail to identify local or regional progressive thinning of the

inner macular layers. Hence, estimating macular thickness changes in smaller areas such as $3^\circ \times 3^\circ$ superpixels is of interest.^{4,18–23} Miraftebi et al.²⁴ reported that repeated thickness measurements across the macula at the level of $3^\circ \times 3^\circ$ superpixels demonstrated a low and uniform within-session variability. Nouri-Mahdavi et al.²⁵ recently showed that between-session variability of local inner retinal thickness measurements was very low and that most of the total between-session variability could be explained by within-session variability. While measurement of change in localized areas of the macula is essential for detecting local changes, there is a significant correlation between thickness measurements and thickness changes in physically or physiologically adjacent regions of the macula.²⁰ A hierarchical approach, once fully developed, can address the varying behavior of superpixels across the macular region while at the same time taking into account any potential correlations in rates of change or their variance among superpixels of varying sizes.

Several studies have employed linear mixed models for enhancing estimation of functional and structural rates of change in glaucoma eyes.^{26–30} Most such studies addressed functional measurements in glaucoma.^{31–40} Random-effects models can more accurately estimate rates of structural change if appropriately implemented as compared to fixed-effects models that do not take population information into account.^{41–45} Random-effects models are known to be more efficient and accurate in modeling subject-specific effects in longitudinal data, allowing population information to help estimate individual trends.⁴⁶ Random-effects longitudinal models, therefore, would be superior models for analyzing macular thickness measurements over time, both to accommodate within-subject correlations over time and to take advantage of such correlations to more efficiently identify progression.

Bayesian methods have advantages over non-Bayesian methods of estimation in appropriately accommodating all uncertainty in the model and being able to deal with situations where variance components are small without setting the variances to zero. Bayesian approaches are able to handle smaller data sets, estimate subject-specific parameters such as slopes, and provide appropriate measures of uncertainty.

The aim of this study is to validate a hierarchical model for estimating rates of change of macular ganglion cell complex (GCC) measurements at the level of superpixels in a cohort of glaucoma eyes followed over a period of 4 years and to determine the relationship between baseline GCC thickness measurements

and corresponding local rates of change for the population and for individuals.

Methods

Study Sample

In total, 112 eyes from the Advanced Glaucoma Progression Study, an ongoing longitudinal study at the University of California, Los Angeles (UCLA), with a minimum of four macular OCT images, ≥ 2 years of follow-up, and no other ocular pathology at baseline and during follow-up were recruited. We analyzed observations up to 4.2 years after baseline in this study. All data from visits less than 0.2 years after a previous visit were omitted.

The study adhered to the tenets of the Declaration of Helsinki, was approved by UCLA's Human Research Protection Program, and conformed to the Health Insurance Portability and Accountability Act policies. All patients provided written informed consent at the time of enrollment in the study.

The enrolled eyes met the following inclusion criteria: (a) clinical diagnosis of primary open-angle glaucoma, pseudoexfoliative glaucoma, pigmentary glaucoma, or primary angle-closure glaucoma and (b) evidence of either central damage on 24-2 visual field (VF), defined as two or more points within the central 10° with $P < 0.05$ on the pattern deviation plot or VF mean deviation (MD) worse than -6 dB. Exclusion criteria were baseline age less than 40 years or greater than 80 years, best-corrected visual acuity $< 20/50$, refractive error exceeding 8 diopters (D) of sphere or 3 D of cylinder, or any significant retinal or neurologic disease potentially affecting OCT measurements.

Imaging Procedures

Macular imaging was carried out with the Posterior Pole Algorithm of Spectralis spectral domain OCT (Heidelberg Engineering, Heidelberg, Germany). The Posterior Pole Algorithm acquires $30^\circ \times 25^\circ$ volume scans of the macula (61 B-scans spaced approximately $120 \mu\text{m}$) centered on the fovea. The software segments the central $24^\circ \times 24^\circ$ of the measurement cube and presents data in an 8×8 array of $3^\circ \times 3^\circ$ superpixels. Each B-scan was repeated 9 to 11 times to reduce speckle noise.

Automated segmentation of individual retinal layers was performed with the Glaucoma Module Premium Edition software (Heidelberg Engineering, Heidelberg, Germany) before data export. Images were reviewed for segmentation errors and image artifacts. Any

1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8
2.1	2.2	2.3	2.4	2.5	2.6	2.7	2.8
3.1	3.2	3.3	3.4	3.5	3.6	3.7	3.8
4.1	4.2	4.3	4.4	4.5	4.6	4.7	4.8
5.1	5.2	5.3	5.4	5.5	5.6	5.7	5.8
6.1	6.2	6.3	6.4	6.5	6.6	6.7	6.8
7.1	7.2	7.3	7.4	7.5	7.6	7.7	7.8
8.1	8.2	8.3	8.4	8.5	8.6	8.7	8.8

Figure 1. The 8 × 8 grid of 64 superpixels from the macular volume scan as provided by Spectralis OCT’s Posterior Pole Algorithm. The central 36 superpixels (shown in gray) were selected for final analyses.

segmentation errors the reviewer believed could be improved by manual correction were manually corrected with the OCT device’s built-in software. The 8 × 8 arrays for retinal nerve fiber layer, ganglion cell layer, and inner plexiform layer (IPL) were summed up to create the GCC measurements, which represents data delimited by the internal limiting membrane and the IPL/inner nuclear layer boundary. Due to significant variability observed in the peripheral regions of the macular volume scans, we chose to include only the central 36 (6 × 6 or 18° × 18°) superpixels for further analyses (Fig. 1).

Data Inspection, Exploration, and Outlier Removal

A single observation y_{ijk} is the GCC thickness observed on subject i at time t_{ij} at subject i ’s j th visit, in superpixel k , where the time of the first visit for all subjects is $t_{ij} = 0$. A profile is the time-ordered sequence of observations (t_{ij}, y_{ijk}) from $j = 1$ to n_{ik} , where n_{ik} is the number of observations on superpixel k for subject i . Profiles were plotted by connecting consecutive observations with line segments. Plotting data helps in identifying appropriate classes of models for fitting data. We plotted profiles for all subjects and superpixels to look for outliers, check for the time trend, and understand variability within a person and between people. Profile plots suggested that outliers were present.

Outlier Detection and Removal

Because of the large number of observations, we developed a semiautomated process to identify outliers. For each person and superpixel, we calculated consecutive visit absolute differences $|y_{ijk} - y_{i(j-1)k}|$ and consecutive visit absolute centered slopes, $|(y_{ijk} - y_{i(j-1)k}) / (t_{ij} - t_{i(j-1)}) + 0.5|$. Slopes were centered around $-0.5 \mu\text{m}/\text{year}$ as that was the mean of the pooled set of slopes across all subjects, superpixels, and pairs of consecutive visits. We plotted histograms of the differences and slopes and examined the largest values. Large absolute differences and large absolute centered slopes are caused by unusual GCC measurements that are high or low compared to preceding or subsequent visit measurements. We flagged absolute slopes greater than $24 \mu\text{m}/\text{year}$ with differences greater than $5 \mu\text{m}$ between sequential visits; this identified the ij th and $i(j-1)$ th observations as candidates for removal. We calculated the sum of absolute visit differences in each superpixel for each subject,

$$\sum_{j=2}^{n_{ik}} |y_{ij} - y_{i(j-1)}|$$

and deleted the candidate outlying observation that caused the largest decrease in the sum of absolute visit differences. If we deleted an observation, we then treated the reduced data set with the same algorithm to see if another observation from the same profile should be deleted as well. Because visits less than 0.2 years after a previous visit generated a large number of candidate outliers, we opted instead to remove all short-term successor visits less than 0.2 years following a previous visit.

Empirical Residuals

We calculated empirical residuals as $e_{ijk} = y_{ijk} - \bar{y}_{ik}$, where $\bar{y}_{ik} = \sum y_{ijk} / n_{ik}$, and the summation is over all n_{ik} observations for subject i in superpixel k . Using empirical residuals in profile plots makes it easier to understand time trends as compared to plotting raw data profiles.⁴⁵ We calculated averages and standard errors of the empirical residuals e_{ijk} in 6-month windows and plotted the averages ± 2 standard errors against time to understand the population time trend for each superpixel.

Hierarchical Longitudinal Models and Inferences

We modeled data from each superpixel k separately with a set of hierarchical random-effects models. Our initially preferred model was linear in time, with a linear population time trend $\alpha_{0k} + \alpha_{1k}t_{ij}$ and random

intercepts β_{0ik} and random slopes β_{1ik} (RIAS),

$$y_{ijk} = \alpha_{0k} + \alpha_{1k}t_{ij} + \beta_{0ik} + \beta_{1ik}t_{ij} + \varepsilon_{ijk} \quad (1)$$

where for superpixel k ,

- α_{0k} is the population intercept, the average intercept at time $t_{ij} = 0$;
- α_{1k} is the population average slope;
- β_{0ik} is the i th subject's random intercept: the unknown difference between subject i 's intercept and the population intercept α_{0k} ;
- β_{1ik} is the i th subject's random slope: the unknown difference between subject i 's slope and α_{1k} ; and
- ε_{ijk} is a random error at time t_{ij} .

While model (1) was our preferred model prior to fitting to data, we considered several variations. We considered a fixed and random intercept (RI)–only model where subject GCC profiles are constant over time; that is, in Equation (1), we set $\alpha_{1k} = \beta_{1ik} = 0$; a random intercept and fixed slope (RIFS) model, where $\beta_{1ik} = 0$ but α_{1k} is free to vary; and a random intercept and slope and fixed quadratic (RIASFQ) model, which adds a term $\alpha_{2k}t_{ij}^2$ to the right-hand side of Equation (1). This is a total of four models for the population and individual time trends: RI, RIFS, RIAS, and RIASFQ. We modeled residuals ε_{ijk} as normal with unknown variance σ_k^2 ,

$$\varepsilon_{ijk} | \sigma_k^2 \sim N(0, \sigma_k^2).$$

Alternatively, we considered a random-effects model where each eye has its own residual variance σ_{ik}^2 ,

$$\varepsilon_{ijk} | \sigma_{ik}^2 \sim N(0, \sigma_{ik}^2).$$

We also considered the possibility of autocorrelation between consecutive observations within an eye and superpixel with fixed σ_k^2 or random σ_{ik}^2 residual variances. Thus, we considered $4 * 2 * 2 = 16$ distinct models.

We modeled the random effects, $\beta_{ik} = \beta_{i0k}$ in the RI models or $\beta_{ik} = (\beta_{i0k}, \beta_{i1k})'$ in the RIAS models, as normal $\beta_{ik} | D \sim N(0, D_k)$, where D_k is scalar for the RI models and D_k is a 2×2 matrix for the RIAS models:

$$D_{k2 \times 2} = \begin{pmatrix} D_{k11} & D_{k12} \\ D_{k21} & D_{k22} \end{pmatrix}$$

We are particularly interested in the correlation

$$\rho_k = \frac{D_{k12}}{(D_{k11} D_{k22})^{1/2}}$$

between the intercepts and slopes. Priors for σ_{ik} were inverse gamma with unknown mean σ_{mk} and

unknown standard deviation σ_{sk} . For the hyperparameters α_{0k} , α_{1k} , σ^2 , D_k , ρ_k , and $(\sigma_{mk}, \sigma_{sk})$, we used convenient prior distributions. Parameters of these distributions were specified so that they worked well and were appropriate for all superpixels. While this approach uses the cohort data to specify these priors, data from each superpixel represent only 1/36th of the entire superpixel data, so very little of each superpixel data contributed to the prior construction. Full model specifications and exact parameter values of the priors are given in the online supplement. We plotted summary measures of our analyses over the 6×6 grid of superpixels.

For each superpixel, we calculated one-sided Bayesian P values defined as the posterior probability that the population slope is greater than zero. This is a measure of how well the sign of the population slope is determined and corresponds asymptotically to a one-sided frequentist P value. This number can be doubled to construct a two-sided P value but then loses its Bayesian interpretation. Bayesian P values that are near zero or 1 indicate a slope whose sign is well determined; we used $P < 0.025$ and $P > 0.975$ as traditional cutoffs to identify statistical significance. We discuss this in more detail in the online supplement.

We inspected posterior distributions of individual slopes $\alpha_{1k} + \beta_{1ik}$, and we calculated the posterior probability $p_{ik} = P(\alpha_{1k} + \beta_{1ik} > 0 | Y)$ that the slopes were positive. Small values of $p_{ik} < 0.025$ identify subjects whose GCC thickness is decreasing at a significant rate over time. We report the fraction of eyes that are decreasing in thickness, determined as the fraction of eyes whose slopes $\alpha_{1k} + \beta_{1ik}$ are significantly negative or significantly positive.

Model Comparison, Model Checks, and Computations

We compared models using the Watanabe–Akaike information criterion (WAIC), with smaller WAIC values representing a better model.^{47,48} We also explored an alternative criterion called the widely applicable Bayesian information criterion (WBIC) and compared the results to WAIC for model selection. WBIC led to similar results as WAIC, and results for WBIC are provided in the online supplement.

For each superpixel, we inspected profile plots of the data and plots of the posterior mean of the residuals $\bar{\varepsilon}_{ijk} = Y_{ijk} - E[\alpha_{0k} + \alpha_{1k}t_{ij} + \beta_{0ik} + \beta_{1ik}t_{ij} | Y_k]$, where Y_k represents all the data for the k th superpixel.⁴⁵

Computations were run on the R platform using JAGS for the Bayesian computations using UCLA's Hoffman Cluster.^{49,50} We ran three chains with 50,000

iterations of burn-in followed by 200,000 samples with a thin of 10, resulting in 60,000 samples from each posterior.

Results

Data and Outliers

We started with 29,628 observations in 36 superpixels of 112 unique eyes (112 patients) from 823 visits. Inspection of profile plots showed occasional outliers; Supplementary Figure S1 plots example profiles with outliers that are to be removed. We identified and removed 145 (0.5%) outlying data points; therefore, 29,483 observations contributed to subsequent analyses. Table 1 summarizes the clinical and demographic characteristics of the study patients. The mean (SD) age and median (interquartile range) of baseline 24-2 visual field mean deviation (MD) were 66.9 (8.4) years and -6.8 (-12.2 to -4.3) dB, respectively. The average follow-up time was 3.60 (0.44) years with a median (range) of 3.63 (1.94–4.20) years.

We reviewed the macular volume scans of 40 eyes in whom outliers were noted based on the criteria above and identified the following reasons for presence of outliers: one eye had local inner retinal retinoschisis corresponding to superpixels that were marked as outliers, one eye was found to have a mild epiretinal membrane in the superpixels of interest, four eyes had thick vessels that affected thickness measurements in the corresponding areas, and the remaining superpixels displayed poor-quality single B-scans, in which

manual correction could not provide accurate thickness measurements.

Example profile plots after removing outliers are presented in Supplementary Figure S2 (showing profiles for all 112 subjects at superpixel 5.5) and Supplementary Figure S3 (showing profiles for all 36 superpixels for 5 random subjects).

Figure 2 presents summary plots of the empirical residuals for each of the central 36 superpixels. Many superpixels showed a strong decreasing linear population time trend, with superpixels 4.3 to 4.7, 5.3 to 5.7, and 6.4 to 6.7 demonstrating the largest negative slopes; the remaining superpixels had slight negative slopes or were flat.

Model Selection

Among the 16 models we considered, models with autocorrelation were rarely better based on WAIC than corresponding models without autocorrelation and were never the best model, and we do not consider those further. Table 2 presents WAIC values for the remaining 8 models for all 36 superpixels. Smaller WAIC indicates a better model. Random residual variance models were always substantially better than the fixed residual variance models. Consistently, the best models were random intercepts and slopes with random residual variances and either a linear population time trend or a quadratic population time trend. Given that we observed strong linear trends rather than strong population quadratic trends (Fig. 2), we felt that a linear population time trend was more justifiable and appropriate and for the remainder of this article, we discuss only the random intercept and slope model with random residual variances. Additionally, the longitudinal empirical summary plots of standardized residuals in Supplementary Figure S4 did not suggest strong quadratic trends.

Supplementary Table S1 (web appendix) presents summaries of the posteriors of the parameters for all 36 superpixels for the random intercept and slope and random variance model. Figure 3 provides grayscale plots of the posterior means of the population intercepts, population slopes, Bayesian P value of the population slopes, the posterior mean of the correlation between slopes and intercepts and Bayesian P value of the correlation, and the proportion of individual subject slopes that were significantly negative. The largest population slopes were observed in superpixels 4.3 to 4.7, 5.6, 5.7, 6.5 and 6.6. Superpixels whose posterior mean population slope was below -0.2 microns per year (meaning decreasing more rapidly than 0.2 microns per year) had slopes that were significantly negative (one-sided Bayesian $P < 0.01$

Table 1. Demographic and Clinical Characteristics of the Study Eyes

Characteristic	Value
Age, mean (SD), y	66.9 (8.5)
Gender, No. (%)	
Female	70 (62.5)
Male	42 (37.5)
Ethnicity, No. (%)	
White	59 (52.7)
Asian	24 (21.4)
African American	15 (13.4)
Hispanic	14 (12.5)
Baseline 10-2 MD (dB), mean (SD)	-8.9 (5.9)
Baseline 24-2 MD (dB), mean (SD)	-8.7 (6.1)
Follow-up, mean (SD), y	3.60 (0.44)
Baseline GCC (μm), mean (SD)	77 (20.2)

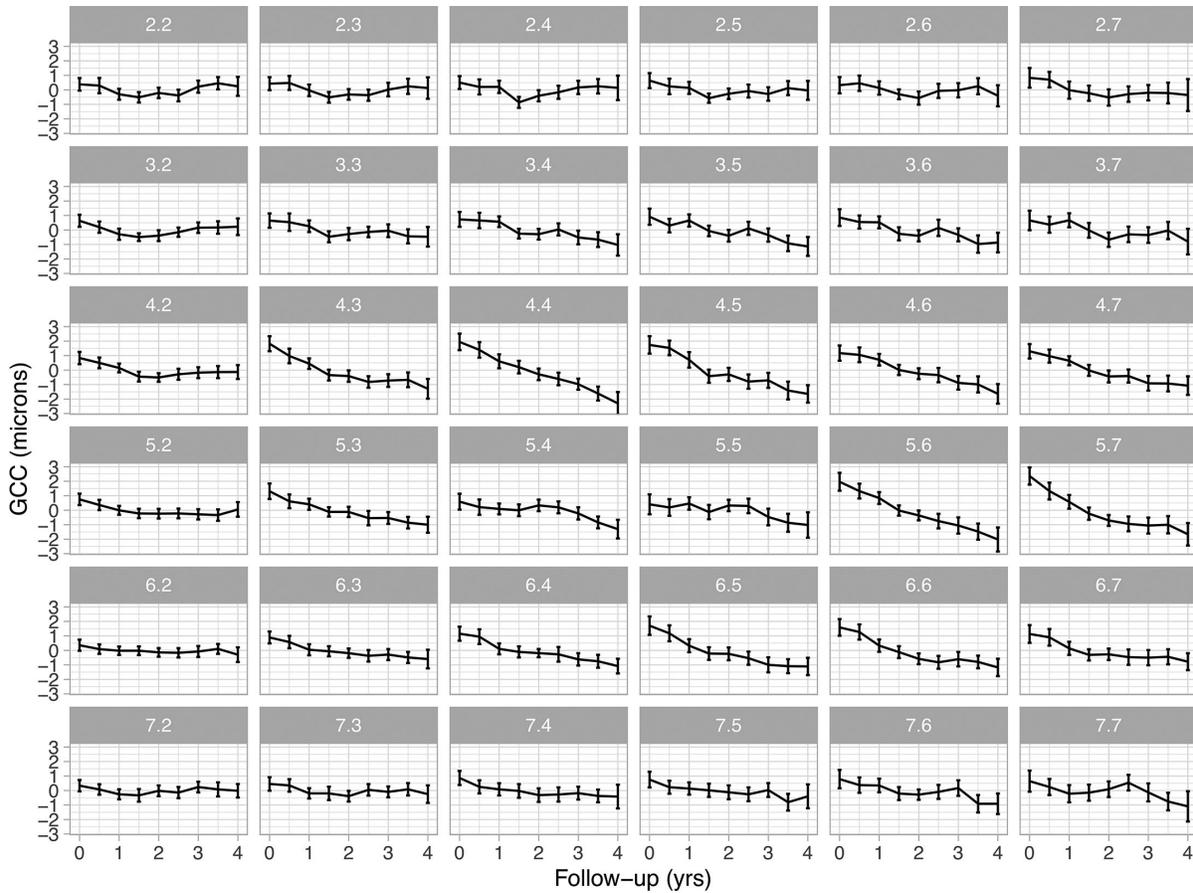


Figure 2. Empirical summary plots of the average trend of mean centered ganglion cell complex thickness across all 36 macular superpixels.

and mostly $P < 0.001$). The fastest thinning (most negative) population superpixels were mostly located in the paracentral and nasal (papillomacular) region. The parafoveal superpixel population slopes were most likely to be significantly less than zero. Point estimates of the correlations between random intercepts and random slopes were negative in 34 of 36 superpixels and were significantly negative in 20 of the 36 superpixels; negative correlations stronger than -0.25 were always significant and correlations weaker than -0.22 were never significant. Superpixels in the super-temporal and inferior regions tended to demonstrate the highest correlation between the baseline thickness (the intercept from the hierarchical model) and their rates of change. The top five coefficients for the correlation between random slopes and baseline thickness at the level of superpixels ranged from -0.43 to -0.50 .

The proportion of individual significantly negative slopes at Bayesian $P < 0.025$ are given in the last plot of Figure 3. Superpixels 4.3 to 4.6, 5.6 to 5.7, and 6.5 to 6.6 had at least 30% of individual eyes that were identified as significantly decreasing (Bayesian $P < 0.025$).

Discussion

We present the results of Bayesian hierarchical models applied to longitudinal and multidimensional macular GCC thickness measurements at the level of $3^\circ \times 3^\circ$ superpixels. This is the first study investigating this type of data within the proposed framework. We designed rules for excluding outliers, reviewed longitudinal trends at the superpixel level, investigated the possibilities of a fixed quadratic component as well as autocorrelation over time in addition to linear fixed and random intercepts and slopes, and estimated the correlation between individual baseline thickness measurements and rates of change.

Our final algorithm for removal of outliers excluded 145 observations representing 0.5% of the total number of available superpixels. The removed outliers were all visually distinctive based on inspection of profile plots. We further checked patients' raw B-scans to detect issues and found many problems with the removed observations. Not all outliers may be identified by any statistical algorithm, including the one used in this

Table 2. WAIC for Each of the Eight Models (Without Autocorrelation) Fit to Data From Each Superpixel

Superpixel	Fixed Variance Models				Random Variance Models			
	RI	RIFS	RIRS	RIRSFQ	RI	RIFS	RIRS	RIRSFQ
2.2	3774.80	3776.92	3621.54	3605.51	3605.24	3607.89	3510.99	3481.50
2.3	3832.67	3832.35	3716.20	3705.80	3683.18	3684.41	3614.65	3601.72
2.4	3896.34	3897.93	3764.03	3757.12	3747.50	3749.68	3655.79	3647.36
2.5	3857.61	3853.54	3696.44	3689.04	3661.43	3660.54	3561.21	3548.37
2.6	4009.56	4007.89	3871.53	3872.16	3798.94	3798.70	3743.39	3743.14
2.7	4303.81	4291.94	4038.73	4035.85	3956.03	3958.96	3835.46	3832.98
3.2	3600.96	3601.57	3411.11	3384.68	3467.08	3469.26	3331.76	3295.26
3.3	3915.21	3898.01	3676.32	3677.27	3644.17	3637.91	3557.03	3557.47
3.4	3920.41	3875.03	3584.25	3586.11	3685.31	3665.66	3542.33	3543.87
3.5	3982.34	3932.00	3666.24	3665.38	3769.80	3735.89	3601.25	3602.82
3.6	4076.30	4036.22	3736.94	3738.35	3783.85	3744.33	3609.05	3610.57
3.7	4254.84	4238.02	3974.57	3975.13	3967.64	3960.06	3803.62	3805.85
4.2	3588.15	3564.81	3339.35	3322.81	3463.24	3455.51	3279.29	3267.98
4.3	3996.82	3870.40	3594.75	3572.00	3718.20	3626.27	3449.64	3429.23
4.4	4157.36	3949.33	3779.14	3781.42	4000.50	3825.67	3731.80	3733.80
4.5	4258.22	4116.84	3911.79	3911.44	4141.84	4013.98	3874.16	3874.95
4.6	4036.47	3926.13	3525.89	3526.47	3730.53	3628.01	3367.98	3365.32
4.7	3981.74	3874.25	3631.33	3632.80	3729.78	3612.14	3476.21	3479.19
5.2	3496.77	3475.88	3329.95	3317.88	3381.16	3374.06	3259.92	3245.47
5.3	3914.25	3830.04	3545.99	3547.07	3554.34	3492.85	3403.12	3405.10
5.4	3911.69	3885.50	3654.75	3646.99	3663.38	3660.93	3507.41	3489.10
5.5	4218.13	4203.12	3791.65	3783.09	4009.03	3986.32	3742.10	3731.46
5.6	4220.45	4031.34	3606.49	3607.11	3958.04	3803.42	3514.32	3515.53
5.7	4278.88	4119.37	3862.60	3839.79	4033.10	3859.94	3698.80	3664.28
6.2	3402.13	3400.46	3256.24	3255.20	3322.04	3322.06	3199.52	3194.18
6.3	3685.76	3643.89	3488.99	3486.58	3512.02	3506.20	3404.90	3402.78
6.4	3888.18	3811.88	3492.60	3494.41	3642.45	3582.33	3431.45	3429.70
6.5	4177.54	4066.24	3748.77	3745.82	3833.85	3747.39	3549.54	3548.15
6.6	4055.69	3946.21	3690.24	3683.41	3749.28	3652.55	3530.10	3521.71
6.7	4107.33	4064.23	3783.13	3781.77	3851.76	3811.73	3668.85	3663.21
7.2	3645.33	3647.61	3608.10	3601.47	3484.74	3485.43	3451.64	3422.26
7.3	3730.56	3729.19	3645.24	3639.22	3571.06	3567.10	3462.11	3444.69
7.4	3939.05	3921.83	3819.27	3817.01	3745.34	3728.61	3642.23	3632.47
7.5	4001.91	3984.82	3843.98	3846.21	3723.73	3708.95	3601.86	3600.90
7.6	4094.93	4071.44	3790.73	3792.76	3829.88	3820.65	3679.00	3679.97
7.7	4326.06	4317.63	4215.70	4216.41	4137.63	4136.94	4045.78	4047.87
Total	142,538.29	140,693.86	132,714.58	132,541.52	134,557.11	133,082.33	128,338.28	128,060.22

For each superpixel and for the four models with constant residual variance, the best model WAIC is bolded and similarly for the four models with random residual variance. The best model across all eight models is shaded for each superpixel.

study; we believe that the outliers we identified were mostly incorrect or inappropriate measurements. A lower threshold for detection of outliers would result in excluding potentially valid data. Therefore, we kept as much as possible of valid data to provide accurate results from the model. We recently reported that GCC is potentially the optimal inner macular parameter for

detection of glaucoma progression across the spectrum of glaucoma severity,^{4,20} and thus we used GCC as the macular outcome of interest for this study.

We believe that the proposed model is appropriate for estimating rates of change of GCC thickness measures over time for several reasons. First, to properly estimate slopes and their uncertainty, a correct

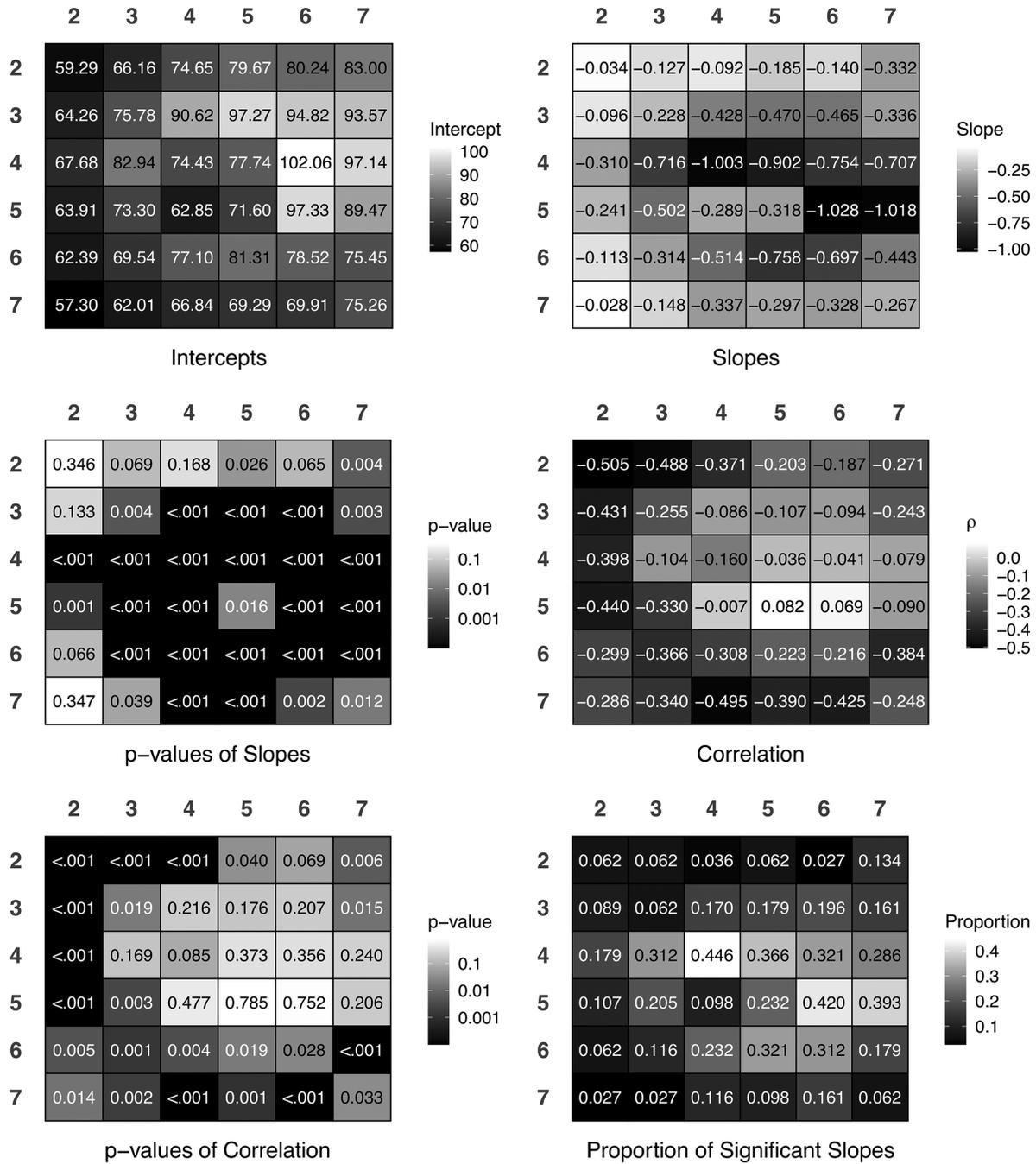


Figure 3. Summary of the results from the best-fit hierarchical model (random intercept and random slope with random variance). The grayscale grids display the distribution of (A) the estimated intercepts, (B) slopes, (C) *P* values for slopes, (D) correlation between random intercepts and random slopes, (E) *P* values for correlation, and (F) the proportion of subjects with significant slopes. The grids show the central 36 superpixels (out of 64 superpixels provided by Spectralis OCT’s Posterior Pole Algorithm) selected for fitting hierarchical models. The key to the grayscale maps is provided on the right side of each image.

or near-correct covariance structure has to be identified for within-eye residuals for the GCC thickness measurements. Based on our results, the model needs to allow for different residual variances for different eyes and random subject-specific intercepts and slopes. We also investigated the possibility of an autoregressive

variance structure for the within-eye measurements; this would mean that residuals for individual eyes are correlated in time even after accounting for random intercepts and slopes. We found no evidence for an autoregressive variance-covariance structure in the final model.

The final model presented here is a hierarchical linear random intercept and random slope model. This model allows for individual superpixels to start at different baseline levels and to have varying slopes within and between individuals. A model incorporating a fixed quadratic effect performed as well as or better than the linear model based on WAIC, but the quadratic trend was not reflected in the profile plots (e.g., see Supplementary Figures S2 and S3), empirical summary plot (Fig. 2), or Supplementary Figure S4, which is similar to the empirical summary plot but is the empirical summary plot for the standardized residuals of our preferred model. There was some evidence for population-level short-term nonlinearity over time, but we are unable to formulate rationales for nonlinear population trends that later return to linearity.

Overall, the magnitude of global perimetric change was fairly small in this cohort (about 1-dB change in the 10-2 visual field MD). However, study of localized structural rates at the level of $3^\circ \times 3^\circ$ superpixels and our Bayesian model allowed detection of significant local or regional trends both in the population and for individual eyes. Linear regression applied to measurements from single eyes will not have the same power to detect trends in individual eyes. Linear regression that treats observations over time within an eye as independent and is applied to longitudinal series from multiple eyes is inappropriate for identifying population trends and will vastly overstate the significance of the resulting inferences.

The central 4 superpixels of the Spectralis macular volume scan are unique because the RGC density in the perifoveal region changes drastically, rising from near zero to near the peak level of density reached in the outer adjacent superpixels. Smaller superpixels could help in identifying trends more accurately in the perifoveal region. The rates of change at the superpixel level tended to be most rapid (most negative) in the perifoveal area and the papillomacular bundle (Fig. 3), but the strongest correlations between the baseline thickness and rates of change were actually seen in the temporal macula.

We were able to identify 75% (27) of the 36 superpixels as having significantly negative population average GCC rates (Bayesian $P < 0.025$) using 4 years of data from 112 subjects. Our Bayesian hierarchical random-effects model is helpful for identifying individual eyes that are thinning at a significant rate because it uses population-level information to help estimate slopes in individual eyes, increasing our chances of identifying individual eyes as decreasing compared to a simple linear regression model analyzing a single individual's observations. The Bayesian approach provides a justified and informative prior for subject-level slopes; the

prior, in turn, yields additional information about a given slope above and beyond the information in an individual's observations.

Bayesian models have been reported to more efficiently identify glaucoma progression when combining various structural and functional measurements.^{16,35,38,40} Our current work represents an initial effort toward developing an optimized hierarchical model for fitting longitudinal structural macular data. Superpixels or even clusters of superpixels across the macula are correlated, with the correlation being a function of physical or physiologic distance between the superpixels.²⁰ We analyzed data for each superpixel separately. This made the modeling easier, at a cost of not allowing us to borrow strength from neighboring superpixels to better estimate trends at each superpixel and preventing us from making joint statements about individual trends across superpixels. Our results are valid despite the lack of a spatial component: posterior means of parameters estimate the corresponding underlying parameters, and standard deviations of the parameters are appropriate summary measures of the uncertainty. For example, posterior estimates, standard deviations, and 95% confidence interval of subject-specific slope estimates could be used as inputs to treatment decisions. Combined spatial-longitudinal models come with high computational costs and have yet to be developed for macular thickness measurements.

We plan future work on multivariate modeling of data from multiple superpixels over time. Statistical models for multivariate macular thickness data need to have longitudinal components, hierarchical components, and spatial components. Each component is a distinct statistical modeling project. Settling on a utile model for the longitudinal component substantially eases the modeling task in the hierarchical and spatial components, and modeling the hierarchical and spatial components would be daunting before an appropriate longitudinal model is identified. Thus, it is now possible to continue to multivariate models incorporating our single most appropriate longitudinal model instead of having to consider multiple possible longitudinal models. Our model has seven hyperparameters for each superpixel: a population intercept and slope (two parameters), random intercept and slope covariance matrix (two variance parameters and one correlation parameter), and two parameters (superpixel mean and variance) for modeling the subject-specific residual variances. The next component modeling step is to incorporate this model in a hierarchical model where each superpixel has a seven-dimensional random effect. The spatial component could then follow, although numerous modeling decisions remain to be explored

for longitudinal structural macular data. For example, a spatial model specification requires a choice of type of spatial correlation model, details of the model specification such as neighborhood structures, and number of spatial components.

Covariates such as age, other demographic variables, and impact of various interventions and other diseases can be considered for predicting macular thickness with both short- and long-term impacts. While the influence of some covariates on macular thickness measurements might be modeled as part of a single research project and a single paper, other covariates such as age and the impact of treatment are likely to result in multiple papers. Effectively, the more interesting the covariate, the harder the modeling project. Identification of regions of interest (i.e., areas of eyes where progression is occurring) requires accurate data and high-quality complex statistical models of which this article is a key first step. Future models will also incorporate functional measurements in joint models for disease deterioration in the central macula. The most crucial RGCs of the human eye reside in this area, and progress in detection of ongoing damage in this region has important uses in the implementation and enhancement of therapeutic measures in patients with glaucoma.

Validation of our model in future data sets is an important task for future work. In the end, our claim is that this model is appropriate for our data set. In validating this model for our data, we compared 16 different models and conducted extensive inspections of our data graphically to eliminate many other possible models. In addition to our conclusion about the best longitudinal model for our data, we have also outlined an approach to modeling future longitudinal data sets of macular thickness measurements over time. Other data sets may include non- or preglaucomatous eyes or may include longer timeframes since disease onset compared to our data set. Similarly, other data sets may include younger patients, different mixes of glaucoma types, and different demographics. These features mean that our model may or may not be appropriate for these other data sets; in contrast, our process is appropriate, and the process can lead to different models for different data sets.

Our results here may speed analyses of new data sets. If initial graphical inspection of the new data set shows decreasing linear time trends, then our model may well be appropriate, and researchers can jumpstart their analyses directly with our model followed by graphical checks of the residuals to confirm the appropriateness. We will make our algorithms publicly available as this body of work is funded by the National Institutes of Health.

Data sets with longer follow-up times or collected on patients with more severe damage at baseline may show floor effects; this would require extensions of our model. Similarly, data sets from younger patients and/or healthy patients may show ceiling effects. We did not see evidence of floor effects in our data despite our extensive graphical data inspection. In particular, Figure 2 did not show population-level floor effects; rather, the population mean continued to decrease at a linear rate throughout the study period. Similarly, Supplementary Figure S2 did not show individual floor effects for superpixel 5.5; we inspected similar plots for all superpixels and observed no floor effects. In the same vein, Supplementary Figure S3 did not show floor effects for the five selected subjects. We reviewed similar plots for all subjects and were unable to ascertain floor effects. In contrast, similar plotting of longitudinal visual field data in patients with glaucoma can and does illustrate floor effects.⁵¹

Over the short term, linearity of change in the macula seems reasonable, as seen in our longitudinal cohort; however, over longer periods of time, nonlinear time trends must be considered and different nonlinear time trends be allowed for different superpixels within an eye and for different eyes. Random errors ε_{ijk} might demonstrate autocorrelation over time, although we did not see that in our analysis. Recent research has suggested that the superpixels in which measurements behave similarly are not just neighbors but may belong to differently shaped nonconvex clusters.²⁰ Similarly, random intercepts, random slopes, and random errors may have complex spatial associations not easily modeled by local spatial autocorrelation models. Our future work will consider spatial location and physiologic relationships of superpixels (clusters) in expanded Bayesian hierarchical models.

It can take years of regular data collection to be able to detect that the GCC thickness of an individual eye is decreasing. Our study includes data collected every 6 months for 4 years. However, practitioners may not want or need to wait until enough data have been collected and small values of p_{ik} have been reached before deciding on a change in treatment. Rather, one might use a more generous cutoff of, for example, 0.1 instead of 0.025, for flagging a clinically significant thinning of the GCC in an individual eye.

Translational Relevance

We provide an initial Bayesian hierarchical model for estimating local rates of change for GCC thickness measurements across the macula. Such estimates are better estimated by using data from the entire patient sample to estimate individual eye superpixel rates

of change and appropriately modeling the correlation between repeated measurements over time. The proposed Bayesian hierarchical model can be applied to various macular outcome measures from different OCT devices and to superpixels of variable sizes to estimate local rates of change and progression probability. Future models will have to be able to handle the very complex anatomic and physiologic correlations between macular superpixels. Such models, when optimized, will provide clinicians with estimates of rates of change and the probability of progression globally and at the level of superpixels or clusters of superpixels at any given point in time and also provide prediction probabilities for the future course of the disease. Structural and functional rates of change may be combined in joint hierarchical models to better define the longitudinal structure–function relationships in the macular region in glaucoma.

Acknowledgments

Supported by National Institutes of Health R01 grant (R01-EY029792), an unrestricted departmental grant from Research to Prevent Blindness, and an unrestricted grant from Heidelberg Engineering.

Disclosure: **V. Mohammadzadeh**, None; **E. Su**, None; **S. Heydar Zadeh**, None; **S.K. Law**, None; **A.L. Coleman**, None; **J. Caprioli**, None; **R.E. Weiss**, None; **K. Nouri-Mahdavi**, Heidelberg Engineering

References

1. Quigley HA, Dunkelberger GR, Green WR. Retinal ganglion cell atrophy correlated with automated perimetry in human eyes with glaucoma. *Am J Ophthalmol*. 1989;107(5):453–464.
2. Bowd C, Zangwill LM, Weinreb RN, Medeiros FA, Belghith A. Estimating optical coherence tomography structural measurement floors to improve detection of progression in advanced glaucoma. *Am J Ophthalmol*. 2017;175:37–44.
3. Nguyen AT, Greenfield DS, Bhakta AS, Lee J, Feuer WJ. Detecting glaucoma progression using guided progression analysis with OCT and visual field assessment in eyes classified by International Classification of Disease severity codes. *Ophthalmol Glaucoma*. 2019;2(1):36–46.
4. Mohammadzadeh V, Rabiolo A, Fu Q, et al. Longitudinal macular structure-function relationships in glaucoma. *Ophthalmology*. 2020;127(7):888–900.
5. Results of the European Glaucoma Prevention Study. *Ophthalmology*. 2005;112(3):366–375.
6. Harwerth RS, Carter-Dawson L, Shen F, Smith EL, III, Crawford ML. Ganglion cell losses underlying visual field defects from experimental glaucoma. *Invest Ophthalmol Vis Sci*. 1999;40(10):2242–2250.
7. Sommer A, Katz J, Quigley HA, et al. Clinically detectable nerve fiber atrophy precedes the onset of glaucomatous field loss. *Arch Ophthalmol*. 1991;109(1):77–83.
8. Artes PH, Chauhan BC. Longitudinal changes in the visual field and optic disc in glaucoma. *Prog Retin Eye Res*. 2005;24(3):333–354.
9. Mohammadzadeh V, Fatehi N, Yarmohammadi A, et al. Macular imaging with optical coherence tomography in glaucoma. *Surv Ophthalmol*. 2020;65(6):597–638.
10. Lee KS, Lee JR, Na JH, Kook MS. Usefulness of macular thickness derived from spectral-domain optical coherence tomography in the detection of glaucoma progression. *Invest Ophthalmol Vis Sci*. 2013;54(3):1941–1949.
11. Naghizadeh F, Garas A, Vargha P, Hollo G. Detection of early glaucomatous progression with different parameters of the RTVue optical coherence tomograph. *J Glaucoma*. 2014;23(4):195–198.
12. Nieves-Moreno M, Martinez-de-la-Casa JM, Bambo MP, et al. New normative database of inner macular layer thickness measured by Spectralis OCT used as reference standard for glaucoma detection. *Transl Vis Sci Technol*. 2018;7(1):20.
13. Chien JL, Ghassibi MP, Patthanathamrongkasem T, et al. Glaucoma diagnostic capability of global and regional measurements of isolated ganglion cell layer and inner plexiform layer. *J Glaucoma*. 2017;26(3):208–215.
14. Martucci A, Toschi N, Cesareo M, et al. Spectral domain optical coherence tomography assessment of macular and optic nerve alterations in patients with glaucoma and correlation with visual field index. *J Ophthalmol*. 2018;2018:6581846.
15. Na JH, Sung KR, Baek S, et al. Detection of glaucoma progression by assessment of segmented macular thickness data obtained using spectral domain optical coherence tomography. *Invest Ophthalmol Vis Sci*. 2012;53(7):3817–3826.
16. Tatham AJ, Medeiros FA. Detecting structural progression in glaucoma with optical coherence tomography. *Ophthalmology*. 2017;124(12S):S57–S65.
17. Sung KR, Sun JH, Na JH, Lee JY, Lee Y. Progression detection capability of macular thickness

- in advanced glaucomatous eyes. *Ophthalmology*. 2012;119(2):308–313.
18. Miraftebi A, Amini N, Morales E, et al. Macular SD-OCT outcome measures: comparison of local structure-function relationships and dynamic range. *Invest Ophthalmol Vis Sci*. 2016;57(11):4815–4823.
 19. Lee JW, Morales E, Sharifipour F, et al. The relationship between central visual field sensitivity and macular ganglion cell/inner plexiform layer thickness in glaucoma. *Br J Ophthalmol*. 2017;101(8):1052–1058.
 20. Rabiolo A, Mohammadzadeh V, Fatehi N, et al. Comparison of rates of progression of macular OCT measures in glaucoma. *Transl Vis Sci Technol*. 2020;9(7):50.
 21. Hood DC, Raza AS, de Moraes CG, Liebmann JM, Ritch R. Glaucomatous damage of the macula. *Prog Retin Eye Res*. 2013;32:1–21.
 22. Hood DC, Xin D, Wang D, et al. A region-of-interest approach for detecting progression of glaucomatous damage with optical coherence tomography. *JAMA Ophthalmol*. 2015;133(12):1438–1444.
 23. Raza AS, Cho J, de Moraes CG, et al. Retinal ganglion cell layer thickness and local visual field sensitivity in glaucoma. *Arch Ophthalmol*. 2011;129(12):1529–1536.
 24. Miraftebi A, Amini N, Gornbein J, et al. Local variability of macular thickness measurements with SD-OCT and influencing factors. *Transl Vis Sci Technol*. 2016;5(4):5.
 25. Nouri-Mahdavi K, Fatehi N, Caprioli J. Longitudinal macular structure-function relationships in glaucoma and their sources of variability. *Am J Ophthalmol*. 2019;207:18–36.
 26. Jammal AA, Thompson AC, Mariottoni EB, et al. Rates of glaucomatous structural and functional change from a large clinical population: the Duke Glaucoma Registry Study. *Am J Ophthalmol*. 2021;222:238–247.
 27. Chua J, Kadziauskiene A, Wong D, et al. One year structural and functional glaucoma progression after trabeculectomy. *Sci Rep*. 2020;10(1):2808.
 28. Wu Z, Crabb DP, Chauhan BC, Crowston JG, Medeiros FA. Improving the feasibility of glaucoma clinical trials using trend-based visual field progression end points. *Ophthalmol Glaucoma*. 2019;2(2):72–77.
 29. Liebmann K, De Moraes CG, Liebmann JM. Measuring rates of visual field progression in linear versus nonlinear scales: implications for understanding the relationship between baseline damage and target rates of glaucoma progression. *J Glaucoma*. 2017;26(8):721–725.
 30. Zhang C, Tatham AJ, Abe RY, et al. Corneal hysteresis and progressive retinal nerve fiber layer loss in glaucoma. *Am J Ophthalmol*. 2016;166:29–36.
 31. Murata H, Zangwill LM, Fujino Y, et al. Validating variational Bayes linear regression method with multi-central datasets. *Invest Ophthalmol Vis Sci*. 2018;59(5):1897–1904.
 32. Anderson AJ. Estimating the true distribution of visual field progression rates in glaucoma. *Invest Ophthalmol Vis Sci*. 2015;56(3):1603–1608.
 33. Murata H, Araie M, Asaoka R. A new approach to measure visual field progression in glaucoma patients using variational bayes linear regression. *Invest Ophthalmol Vis Sci*. 2014;55(12):8386–8392.
 34. Anderson AJ, Johnson CA. How useful is population data for informing visual field progression rate estimation? *Invest Ophthalmol Vis Sci*. 2013;54(3):2198–2206.
 35. Russell RA, Malik R, Chauhan BC, Crabb DP, DF Garway-Heath. Improved estimates of visual field progression using Bayesian linear regression to integrate structural information in patients with ocular hypertension. *Invest Ophthalmol Vis Sci*. 2012;53(6):2760–2769.
 36. Medeiros FA, Zangwill LM, Weinreb RN. Improved prediction of rates of visual field loss in glaucoma using empirical Bayes estimates of slopes of change. *J Glaucoma*. 2012;21(3):147–154.
 37. Medeiros FA, Zangwill LM, Mansouri K, Lisboa R, Tafreshi A, Weinreb RN. Incorporating risk factors to improve the assessment of rates of glaucomatous progression. *Invest Ophthalmol Vis Sci*. 2012;53(4):2199–2207.
 38. Medeiros FA, Zangwill LM, Girkin CA, Liebmann JM, Weinreb RN. Combining structural and functional measurements to improve estimates of rates of glaucomatous progression. *Am J Ophthalmol*. 2012;153(6):1197–1205.e1191.
 39. Medeiros FA, Weinreb RN, Moore G, Liebmann JM, Girkin CA, Zangwill LM. Integrating event- and trend-based analyses to improve detection of glaucomatous visual field progression. *Ophthalmology*. 2012;119(3):458–467.
 40. Medeiros FA, Leite MT, Zangwill LM, Weinreb RN. Combining structural and functional measurements to improve detection of glaucoma progression using Bayesian hierarchical models. *Invest Ophthalmol Vis Sci*. 2011;52(8):5794–5803.
 41. Ariyo O, Lesaffre E, Verbeke G, Quintero A. Model selection for Bayesian linear mixed models with longitudinal data: sensitivity to the choice of priors. *Commun Stat Simul Comp*. 2019:1–25.

42. Bryan SR, Eilers PHC, Rosmalen JV, et al. Bayesian hierarchical modeling of longitudinal glaucomatous visual fields using a two-stage approach. *Stat Med*. 2017;36(11):1735–1753.
43. Weiss RE. Bayesian methods for data analysis. *Am J Ophthalmol*. 2010;149(2):187–188.e181.
44. Fitzmaurice GM, Laird NM, Ware JH. *Applied longitudinal analysis*. New York, NY: John Wiley & Sons; 2012.
45. Weiss RE. *Modeling Longitudinal Data*. Springer Science & Business Media; 2005.
46. Harville D. Extension of the Gauss-Markov theorem to include the estimation of random effects. *Ann Stat*. 1976;1:384–395.
47. Gelman A, Hwang J, Vehtari A. Understanding predictive information criteria for Bayesian models. *Stat Comput*. 2013;24(6):997–1016.
48. Watanabe S. Asymptotic equivalence of Bayes cross validation and widely applicable information criterion in singular learning theory. *J Machine Learn Res*. 2010;11(12):3571–3594.
49. R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing; 2013, <https://www.R-project.org/>.
50. Denwood MJ. runjags: An R package providing interface utilities, model templates, parallel computing methods and additional distributions for MCMC models in JAGS. *J Stat Soft*. 2016;71(9):1–25.
51. Otarola F, Chen A, Morales E, Yu F, Afifi A, Caprioli J. Course of glaucomatous visual field loss across the entire perimetric range. *JAMA Ophthalmol*. 2016;134(5):496–502.