Supplementary material to 'Diagnosing Glaucoma Progression with Visual Field Data Using a Spatiotemporal Boundary Detection Method'

Samuel I. Berchuck[1], Jean-Claude Mwanza[2] and Joshua L. Warren[3]

[1]Department of Statistical Science and Forge, Duke University
[2]Department of Ophthalmology, University of North Carolina-Chapel Hill
[3]Department of Biostatistics, Yale University

Introduction

This document provides supplementary material to the paper entitled 'Diagnosing Glaucoma Progression with Visual Field Data Using a Spatiotemporal Boundary Detection Method', and is divided into five sections. Section 1 describes the implementation of the Bayesian model proposed in the paper using Markov chain Monte Carlo (MCMC) simulation. Section 2 details the prediction capability of our model, first introduced in Section 4.5 of the main text. Section 3 presents an additional simulation study to assess the impact of mis-specifying ρ and the dissimilarity metric on our results, while Section 4 presents sensitivity analyses to various modeling assumptions in the visual field (VF) analysis of Section 5 from the main text. Section 5 of the online supplementary material includes an additional figure from the analysis.

1 Implementation of the Model

We draw samples from the joint posterior using a MCMC sampler (Metropolis et al., 1953; Geman and Geman, 1984; Gelfand and Smith, 1990). The Tobit model specification is particularly amenable to MCMC, since the latent process has a closed form full conditional distribution. Many of the parameters have conjugate full conditionals. In particular, all of the hyperparameters have closed form solutions, with the exception of ϕ. None of the observational level parameters display obvious conjugacy. However, due to the separable specification of $\boldsymbol{\theta}$, the full conditional for μ_t, $t = 1 \ldots, \nu$ can be written in closed form. Due to the variety of full conditionals, a Gibbs sampler is used with Metropolis steps for parameters lacking conjugacy. For each eye, the MCMC sampler is run for 250,000 iterations after a 10,000 burn-in and thinned to a final sample size of 10,000.

To guarantee that Metropolis acceptance rates are within an acceptable range, pilot adaptation is used (Banerjee et al., 2003). Convergence is assessed through examination of traceplots and the Geweke diagnostic statistic (Geweke, 1992), with no signs of non-convergence observed for any of the 191 VF series. The observed VF data were scaled by 10 along with the visit days being scaled by 365 to improve the stability of the parameter estimation. Finally, we scale our dissimilarity metric by 100 to improve MCMC mixing.

We present the MCMC sampler for the general method. The parameters in the model consist of $[\boldsymbol{\varphi} = (\boldsymbol{\varphi}_1^T, \ldots, \boldsymbol{\varphi}_\nu^T)^T, \boldsymbol{\zeta}, \boldsymbol{\theta}, \boldsymbol{\delta}, \mathbf{T}, \phi]$, and MCMC posterior sampling proceeds in the following manner. The MCMC sampler is implemented in the R package `womblR`.

1. Sample from the full conditional distribution of φ_{it}:

   Begin by sampling φ_{it} from their full conditional distribution for the set $(\{i, t\} : i = 1, \ldots, n_t, t = 1, \ldots, \nu)$. It is clear that each φ_{it} will only depend on the other

observations at time t. The full conditional for φ_{it} can be written as,

$$f\left(\varphi_{it}|Y_{it},\boldsymbol{\zeta},\boldsymbol{\varphi}_{-it},\mu_t,\tau_t^2,\boldsymbol{\alpha}_t\right) \propto f\left(Y_{it}|g^{-1}\left(\varphi_{it}\right),\boldsymbol{\zeta}\right) \times f\left(\varphi_{it}|\boldsymbol{\varphi}_{-it},\mu_t,\tau_t^2,\boldsymbol{\alpha}_t\right).$$

In general, the full conditional distribution has no closed form, requiring the Metropolis algorithm to be used for sampling. However, for VF data we use the Tobit link, $Y_{it} = \max\{0, \varphi_{it}\}$, which yields the following closed form full conditional density,

$$\begin{aligned} f\left(\varphi_{it}|Y_{it},\boldsymbol{\varphi}_{-it},\mu_t,\tau_t^2,\boldsymbol{\alpha}_t\right) &\propto 1(\varphi_{it} \leq 0) \times \mathrm{N}\left(\mathbb{E}_{it},\mathbb{V}_{it}\right) \\ &= \begin{cases} \mathrm{TN}\left(\mathbb{E}_{it},\mathbb{V}_{it},\leq 0\right) & Y_{it} = 0, \\ Y_{it} & Y_{it} > 0, \end{cases} \end{aligned}$$

where $\mathrm{TN}\left(\mu,\sigma^2;\leq 0\right)$ specifies a truncated normal distribution with mean, μ, and variance, σ^2, truncated above by zero. Recall the moments, $\mathbb{E}_{it} = [\rho\sum_{j=1}^{n_t} w_{ij}\left(\boldsymbol{\alpha}_t\right)\varphi_{jt} + (1-\rho)\mu_t]/\Delta$ and $\mathbb{V}_{it} = \tau_t^2/\Delta$ with $\Delta = \rho\sum_{j=1}^{n_t} w_{ij}\left(\boldsymbol{\alpha}_t\right) + (1-\rho)$.

2. Sample from the full conditional distribution for $\boldsymbol{\zeta}$:

$$f(\boldsymbol{\zeta}|\mathbf{Y},\boldsymbol{\varphi},\boldsymbol{\theta},\boldsymbol{\delta},\mathbf{T},\phi) \propto \left[\prod_{t=1}^{\nu}\prod_{i=1}^{n_t} f(Y_{it}|g^{-1}\left(\varphi_{it}\right),\boldsymbol{\zeta})\right] \times f\left(\boldsymbol{\zeta}\right).$$

The full conditional for $\boldsymbol{\zeta}$ depends on the observational model specification. In general, the full conditional distribution for $\boldsymbol{\zeta}$ has no closed form. For VF data there are no nuisance parameters in the likelihood, so it is not necessary to specify this full conditional.

3. Sample from the full conditional distribution of $\boldsymbol{\theta}_{1\cdot}$: (i.e., $\mu_1,\ldots,\mu_\nu$)

It is not immediately obvious, but with some manipulation we can find a closed form full conditional distribution for $\boldsymbol{\theta}_{1\cdot}$. The key is that we can find an analytical form of the distribution $f\left(\boldsymbol{\theta}_{1\cdot}|\boldsymbol{\theta}_{s\cdot},\boldsymbol{\delta},\mathbf{T},\phi\right)$. Using properties of the multivariate normal

distribution we know that $\text{vec}\left(\boldsymbol{\theta}^T\right) \sim \text{MVN}\left(\text{vec}\left(\mathbf{M}^T\right), \mathbf{T} \otimes \boldsymbol{\Sigma}(\phi)\right)$, where $\mathbf{M} = \boldsymbol{\delta}\mathbf{1}_\nu^T$. Define $s = \{2, 3, \ldots, q+2\}$, so that we can write the covariance in block form,

$$\mathbf{T} = \begin{bmatrix} \mathbf{T}_{11} & \mathbf{T}_{1s} \\ \mathbf{T}_{s1} & \mathbf{T}_{ss} \end{bmatrix} \implies \mathbf{T} \otimes \boldsymbol{\Sigma}(\phi) = \begin{bmatrix} \mathbf{T}_{11} \otimes \boldsymbol{\Sigma}(\phi) & \mathbf{T}_{1s} \otimes \boldsymbol{\Sigma}(\phi) \\ \mathbf{T}_{s1} \otimes \boldsymbol{\Sigma}(\phi) & \mathbf{T}_{ss} \otimes \boldsymbol{\Sigma}(\phi) \end{bmatrix}.$$

Using the properties of the multivariate normal distribution we can compute the moments of $f\left(\boldsymbol{\theta}_{1\cdot}|\boldsymbol{\theta}_{s\cdot}, \boldsymbol{\delta}, \mathbf{T}, \phi\right) \sim \text{N}(\mathbb{E}_{1|s}, \mathbb{C}_{1|s})$,

$$\begin{aligned} \mathbb{E}_{1|s} = \mathbb{E}[\boldsymbol{\theta}_{1\cdot}|\boldsymbol{\theta}_{s\cdot}] &= \mathbf{M}_{1\cdot} + \left[\mathbf{T}_{1s} \otimes \boldsymbol{\Sigma}(\phi)\right]\left[\mathbf{T}_{ss} \otimes \boldsymbol{\Sigma}(\phi)\right]^{-1} \text{vec}\left(\boldsymbol{\theta}_{s\cdot} - \mathbf{M}_{s\cdot}\right) \\ &= \mathbf{M}_{1\cdot} + \left[\mathbf{T}_{1s}\mathbf{T}_{ss}^{-1} \otimes \mathbf{I}_\nu\right] \text{vec}\left(\boldsymbol{\theta}_{s\cdot} - \mathbf{M}_{s\cdot}\right), \end{aligned}$$

and

$$\begin{aligned} \mathbb{C}_{1|s} = \mathbb{C}(\boldsymbol{\theta}_{1\cdot}|\boldsymbol{\theta}_{s\cdot}) &= \left[\mathbf{T}_{11} \otimes \boldsymbol{\Sigma}(\phi)\right] - \left[\mathbf{T}_{1s} \otimes \boldsymbol{\Sigma}(\phi)\right]\left[\mathbf{T}_{ss} \otimes \boldsymbol{\Sigma}(\phi)\right]^{-1}\left[\mathbf{T}_{s1} \otimes \boldsymbol{\Sigma}(\phi)\right] \\ &= \left[\mathbf{T}_{1|s} \otimes \boldsymbol{\Sigma}(\phi)\right], \end{aligned}$$

where $\mathbf{T}_{1|s} = (\mathbf{T}_{11} - \mathbf{T}_{1s}\mathbf{T}_{ss}^{-1}\mathbf{T}_{s1})$. Now, we can derive the full conditional for $\boldsymbol{\theta}_{1\cdot}$. We define a matrix specification of the random effects,

$$f\left(\boldsymbol{\varphi}|\boldsymbol{\theta}\right) \sim \text{MVN}\left(\mathbf{Z}_\theta\boldsymbol{\theta}_{1\cdot}, \mathbf{Q}^{-1}\right),$$

where $\mathbf{Z}_\theta = (\mathbf{I}_\nu \otimes \mathbf{1}_M)$ and $\mathbf{Q} = \text{BlockDiag}\left\{\mathbf{Q}\left(\alpha_t\right)/\tau_t^2\right\}$, for $t = 1, \ldots, \nu$. Then, the full conditional is as follows,

$$\begin{aligned} f(\boldsymbol{\theta}_{1\cdot}|\mathbf{Y}, \boldsymbol{\zeta}, \boldsymbol{\varphi}, \boldsymbol{\theta}_{s\cdot}, \boldsymbol{\delta}, \mathbf{T}, \phi) &\propto f(\boldsymbol{\varphi}|\boldsymbol{\theta}) \times f(\boldsymbol{\theta}_{1\cdot}|\boldsymbol{\theta}_{s\cdot}, \boldsymbol{\delta}, \mathbf{T}, \phi) \\ &\propto \exp\left\{ -\frac{1}{2}\Big[\left(\boldsymbol{\varphi} - \mathbf{Z}_\theta\boldsymbol{\theta}_{1\cdot}\right)^T \mathbf{Q}\left(\boldsymbol{\varphi} - \mathbf{Z}_\theta\boldsymbol{\theta}_{1\cdot}\right) \right. \\ &\qquad \left. + \left(\boldsymbol{\theta}_{1\cdot} - \mathbb{E}_{1|s}\right)^T \mathbb{C}_{1|s}^{-1}\left(\boldsymbol{\theta}_{1\cdot} - \mathbb{E}_{1|s}\right) \Big] \right\} \end{aligned}$$

$$\propto \exp\left\{ -\frac{1}{2}\Big[\boldsymbol{\theta}_{1\cdot}\left(\mathbf{Z}_\theta^T\mathbf{Q}\mathbf{Z}_\theta + \mathbb{C}_{1|s}^{-1}\right)\boldsymbol{\theta}_{1\cdot} - 2\boldsymbol{\theta}_{1\cdot}^T\left(\mathbf{Z}_\theta^T\mathbf{Q}\boldsymbol{\varphi} + \mathbb{C}_{1|s}^{-1}\mathbb{E}_{1|s}\right)\Big]\right\}$$

$$\sim \text{MVN}\left(\mathbb{E}_\theta, \mathbb{V}_\theta\right),$$

where, $\mathbb{V}_\theta = \left(\mathbf{Z}_\theta^T\mathbf{Q}\mathbf{Z}_\theta + \mathbb{C}_{1|s}^{-1}\right)^{-1}$ and $\mathbb{E}_\theta = \mathbb{V}_\theta\left(\mathbf{Z}_\theta^T\mathbf{Q}\boldsymbol{\varphi} + \mathbb{C}_{1|s}^{-1}\mathbb{E}_{1|s}\right)$.

4. Sample τ_t^2 using a Metropolis step:

   Each τ_t^2, $t = 1, \ldots, \nu$, is sampled using a Metropolis step given the following quantity that is proportional to the full conditional density,

   $$f\left(\log(\tau_t)|\mathbf{Y}, \boldsymbol{\varphi}, \boldsymbol{\zeta}, \boldsymbol{\theta}_{-\tau_{tk}}, \boldsymbol{\delta}, \mathbf{T}, \phi\right) \propto f\left(\boldsymbol{\varphi}_t|\mu_t, \tau_t^2, \boldsymbol{\alpha}_t\right) \times f\left(\boldsymbol{\theta}_{2\cdot}|\boldsymbol{\theta}_{s\cdot}, \boldsymbol{\delta}, \mathbf{T}, \phi\right),$$

   where $s = \{1, 3, \ldots, q+2\}$.

5. Sample α_{tk} using a Metropolis step:

   Each α_{tk}, $t = 1, \ldots, \nu$ and $k = 1, \ldots, q$, is sampled using a Metropolis step given the following quantity that is proportional to the full conditional density,

   $$f\left(\log(\alpha_{tk})|\mathbf{Y}, \boldsymbol{\varphi}, \boldsymbol{\zeta}, \boldsymbol{\theta}_{-\alpha_{tk}}, \boldsymbol{\delta}, \mathbf{T}, \phi\right) \propto f\left(\boldsymbol{\varphi}_t|\mu_t, \tau_t^2, \boldsymbol{\alpha}_t\right) \times f\left(\boldsymbol{\theta}_{k+2\cdot}|\boldsymbol{\theta}_{s\cdot}, \boldsymbol{\delta}, \mathbf{T}, \phi\right),$$

   where $s = \{1, 2, 3, \ldots, q+2\}$ with k removed.

6. Sample from the full conditional distribution for $\boldsymbol{\delta}$:

   Define $\mathbf{Z}_\delta = \mathbf{1}_\nu \otimes \mathbf{I}_3$ and $\mathbb{C}_{\boldsymbol{\theta}} = \boldsymbol{\Sigma}(\phi) \otimes \mathbf{T}$. Then the full conditional distribution can be derived as follows,

   $$f(\boldsymbol{\delta}|\mathbf{Y}, \boldsymbol{\varphi}, \boldsymbol{\zeta}, \boldsymbol{\theta}, \mathbf{T}, \phi) \propto f(\text{vec}(\boldsymbol{\theta})|\boldsymbol{\delta}, \mathbf{T}, \phi) \times f(\boldsymbol{\delta})$$

$$
\begin{aligned}
&\propto \exp\Bigg\{ -\frac{1}{2}\Big[\left(\text{vec}(\boldsymbol{\theta}) - \mathbf{Z}_\delta\boldsymbol{\delta}\right)^T \mathbb{C}_{\boldsymbol{\theta}}^{-1} \left(\text{vec}(\boldsymbol{\theta}) - \mathbf{Z}_\delta\boldsymbol{\delta}\right) \\
&\qquad\qquad + (\boldsymbol{\delta} - \boldsymbol{\mu}_\delta)^T \boldsymbol{\Omega}_\delta^{-1} (\boldsymbol{\delta} - \boldsymbol{\mu}_\delta) \Big] \Bigg\} \\
&\propto \exp\Bigg\{ -\frac{1}{2}\Big[\boldsymbol{\delta}^T \left(\mathbf{Z}_\delta^T \mathbb{C}_{\boldsymbol{\theta}}^{-1} \mathbf{Z}_\delta + \boldsymbol{\Omega}^{-1} \right) \boldsymbol{\delta} \\
&\qquad\qquad - 2\boldsymbol{\delta}^T \left(\mathbf{Z}_\delta^T \mathbb{C}_{\boldsymbol{\theta}}^{-1} \text{vec}(\boldsymbol{\theta}) + \boldsymbol{\Omega}_\delta^{-1} \boldsymbol{\mu}_\delta \right) \Big] \Bigg\} \\
&\sim \text{N}(\mathbb{E}_\delta, \mathbb{V}_\delta),
\end{aligned}
$$

where, $\mathbb{V}_\delta = \left(\mathbf{Z}_\delta^T \mathbb{C}_{\boldsymbol{\theta}}^{-1} \mathbf{Z}_\delta + \boldsymbol{\Omega}_\delta^{-1} \right)^{-1}$ and $\mathbb{E}_\delta = \mathbb{V}_\delta \left(\mathbf{Z}_\delta^T \mathbb{C}_{\boldsymbol{\theta}}^{-1} \text{vec}(\boldsymbol{\theta}) + \boldsymbol{\Omega}_\delta^{-1} \boldsymbol{\mu}_\delta \right)$.

7. Sample from the full conditional distribution of $\mathbf{T}$:

   We begin by rewriting the quadratic from of $\boldsymbol{\theta}$,

$$
\begin{aligned}
f(\text{vec}(\boldsymbol{\theta}) | \boldsymbol{\delta}, \mathbf{T}, \phi) &\propto \exp\left\{ -\frac{1}{2} \left(\text{vec}(\boldsymbol{\theta}) - \text{vec}(\mathbf{M})\right)^T \left(\boldsymbol{\Sigma}(\phi) \otimes \mathbf{T}\right)^{-1} \left(\text{vec}(\boldsymbol{\theta}) - \text{vec}(\mathbf{M})\right) \right\} \\
&= \exp\left\{ -\frac{1}{2} \text{tr}\left((\boldsymbol{\theta} - \mathbf{M}) \boldsymbol{\Sigma}(\phi)^{-1} (\boldsymbol{\theta} - \mathbf{M})^T \mathbf{T}^{-1} \right) \right\} \\
&= \exp\left\{ -\frac{1}{2} \text{tr}\left(\mathbf{S}_{\boldsymbol{\theta}} \mathbf{T}^{-1} \right) \right\},
\end{aligned}
$$

   where $\mathbf{S}_{\boldsymbol{\theta}} = (\boldsymbol{\theta} - \mathbf{M}) \boldsymbol{\Sigma}(\phi)^{-1} (\boldsymbol{\theta} - \mathbf{M})^T$. Now, the full conditional distribution for $\mathbf{T}$ is straightforward,

$$
\begin{aligned}
f(\mathbf{T} | \mathbf{Y}, \boldsymbol{\varphi}, \boldsymbol{\zeta}, \boldsymbol{\theta}, \boldsymbol{\delta}, \phi) &\propto f(\text{vec}(\boldsymbol{\theta}) | \boldsymbol{\delta}, \mathbf{T}, \phi) \times f(\mathbf{T}) \\
&\propto |\boldsymbol{\Sigma}(\phi) \otimes \mathbf{T}|^{-\frac{1}{2}} \exp\left\{ -\frac{1}{2} \text{tr}\left(\mathbf{S}_{\boldsymbol{\theta}} \mathbf{T}^{-1} \right) \right\} \\
&\quad \times |\mathbf{T}|^{\frac{-(\xi+3+1)}{2}} \exp\left\{ -\frac{1}{2} \text{tr}(\boldsymbol{\Psi} \mathbf{T}^{-1}) \right\} \\
&\propto |\mathbf{T}|^{\frac{-(\xi+\nu+3+1)}{2}} \exp\left\{ -\frac{1}{2} \text{tr}\left([\mathbf{S}_{\boldsymbol{\theta}} + \boldsymbol{\Psi}] \mathbf{T}^{-1} \right) \right\} \\
&\sim IW\left(\xi + \nu, \mathbf{S}_{\boldsymbol{\theta}} + \boldsymbol{\Psi}\right).
\end{aligned}
$$

8. Sample ϕ using a Metropolis step:

   We transform ϕ to the real line to facilitate sampling. Define a new parameter $\Delta = h(\phi) = \log\left(\frac{\phi - a_\phi}{b_\phi - \phi}\right)$, such that $h^{-1}(\Delta) = (b_\phi \exp\{\Delta\} + a_\phi)/(1 + \exp\{\Delta\})$ and $\left|\frac{\partial}{\partial \Delta} h^{-1}(\Delta)\right| \propto \exp\{\Delta\}/(1 + \exp\{\Delta\})^2$. Now we can sample from the transformed proposal distribution, $\Delta^* \sim \mathrm{N}(\Delta, \delta)$, where δ is a tuning parameter. Then we can obtain a proposal of ϕ, $\phi^* = h^{-1}(\Delta^*)$. Finally, if we define $\boldsymbol{R} = \mathrm{vec}(\boldsymbol{\theta}) - \mathbf{Z}_\delta \boldsymbol{\delta}$ the Metropolis ratio can be calculated as

$$
\begin{aligned}
r &= \frac{f(\phi^* | \mathbf{Y}, \boldsymbol{\varphi}, \boldsymbol{\zeta}, \boldsymbol{\theta}, \boldsymbol{\delta}, \mathbf{T})}{f(\phi | \mathbf{Y}, \boldsymbol{\varphi}, \boldsymbol{\zeta}, \boldsymbol{\theta}, \boldsymbol{\delta}, \mathbf{T})} \\
&\propto \frac{f(\boldsymbol{\theta} | \boldsymbol{\delta}, \mathbf{T}, h^{-1}(\Delta^*)) \times \left|\frac{\partial}{\partial \Delta^*} h^{-1}(\Delta^*)\right|}{f(\boldsymbol{\theta} | \boldsymbol{\delta}, \mathbf{T}, h^{-1}(\Delta)) \times \left|\frac{\partial}{\partial \Delta} h^{-1}(\Delta)\right|} \\
&\propto \frac{|\boldsymbol{\Sigma}(\phi^*)|^{-\frac{3}{2}} \exp\left\{-\frac{1}{2}\left[\boldsymbol{R}^T (\boldsymbol{\Sigma}(\phi^*) \otimes \mathbf{T})^{-1} \boldsymbol{R}\right]\right\} \frac{\exp\{\Delta^*\}}{(1+\exp\{\Delta^*\})^2}}{|\boldsymbol{\Sigma}(\phi)|^{-\frac{3}{2}} \exp\left\{-\frac{1}{2}\left[\boldsymbol{R}^T (\boldsymbol{\Sigma}(\phi) \otimes \mathbf{T})^{-1} \boldsymbol{R}\right]\right\} \frac{\exp\{\Delta\}}{(1+\exp\{\Delta\})^2}}.
\end{aligned}
$$

   Now accept ϕ^* with probability $\min\{1, r\}$, otherwise keep ϕ.

9. Repeat steps 1-8 until convergence has been achieved and an adequate number of posterior samples have been obtained post-convergence.

2 Prediction

In Section 4.5 of the main text, we present the posterior predictive distribution for a future set of VF responses as a function of four known densities that are defined as a consequence of our methodology,

$$\begin{aligned} f\left(\mathbf{Y}_{\nu+1}|\mathbf{Y}\right) &= \int_{\Omega} f\left(\mathbf{Y}_{\nu+1}|\boldsymbol{\Omega},\mathbf{Y}\right) f\left(\boldsymbol{\Omega}|\mathbf{Y}\right) d\boldsymbol{\Omega} \\ &= \int_{\Omega} \underbrace{f\left(\mathbf{Y}_{\nu+1}|g^{-1}\left(\boldsymbol{\varphi}_{\nu+1}\right),\boldsymbol{\zeta}\right)}_{1} \underbrace{f\left(\boldsymbol{\varphi}_{\nu+1}|\boldsymbol{\theta}_{\cdot\nu+1}\right)}_{2} \underbrace{f\left(\boldsymbol{\theta}_{\cdot\nu+1}|\boldsymbol{\theta},\boldsymbol{\delta},\mathbf{T},\phi\right)}_{3} \\ &\qquad \times \underbrace{\left(\boldsymbol{\zeta},\boldsymbol{\theta},\boldsymbol{\delta},\mathbf{T},\phi|\mathbf{Y}\right)}_{4} d\boldsymbol{\Omega}. \end{aligned} \tag{1}$$

Equation 1.1 represents the observed likelihood function written in vector form and is problem specific (or in scalar form: $\prod_{i=1}^{n_{\nu+1}} f(Y_{i(\nu+1)}|g^{-1}\left(\varphi_{i(\nu+1)}\right),\boldsymbol{\zeta})$. Equation 1.2 is the joint specification of the random effects. Equation 1.3 is the least trivial of the four distributions to obtain, but using the properties of the conditional multivariate normal distribution in a similar manner to $f\left(\boldsymbol{\theta}_{1\cdot}|\boldsymbol{\theta}_{s\cdot},\boldsymbol{\delta},\mathbf{T},\phi\right)$, the following can be obtained,

$$f\left(\boldsymbol{\theta}_{\cdot\nu+1}|\boldsymbol{\theta},\boldsymbol{\delta},\mathbf{T},\phi\right) \sim \text{MVN}\left(\mathbb{E}_{\boldsymbol{\theta}_{\cdot\nu+1}},\mathbb{C}_{\boldsymbol{\theta}_{\cdot\nu+1}}\right).$$

The moments are $\mathbb{E}_{\boldsymbol{\theta}_{\cdot\nu+1}} = \boldsymbol{\delta} + (\boldsymbol{\Sigma}^{+} \otimes \mathbf{T})(\text{vec}(\boldsymbol{\theta}) - \mathbf{1}_{\nu} \otimes \boldsymbol{\delta})$ and $\mathbb{C}_{\boldsymbol{\theta}_{\cdot\nu+1}} = \boldsymbol{\Sigma}^{*} \otimes \mathbf{T}$ with $\boldsymbol{\Sigma}^{+} = [\boldsymbol{\Sigma}(\phi)]_{\nu+1,1:\nu}[\boldsymbol{\Sigma}(\phi)]^{-1}_{1:\nu,1:\nu}$ and $\boldsymbol{\Sigma}^{*} = [\boldsymbol{\Sigma}(\phi)]_{\nu+1,\nu+1} - \boldsymbol{\Sigma}^{+}[\boldsymbol{\Sigma}(\phi)]_{1:\nu,\nu+1}$. Here $\boldsymbol{\Sigma}(\phi)$ is a temporal correlation matrix including the new time point $\nu+1$, so that $[\boldsymbol{\Sigma}(\phi)]_{\nu+1,1:\nu}$ is a subset including the row $\nu+1$ and columns 1 up to ν. Finally, Equation 1.4 is the posterior distribution obtained in the MCMC sampler from the original model fit.

3 Model Misspecification

In this simulation study, we explore the sensitivity of the modeling results to changing values of ρ and incorrect specifications of the dissimilarity metric. We simulate data from setting D (Section 6 of the main text) with the maximum number of visits (i.e., 21), such that there is temporal correlation and cross-covariance. In this study, we generate data with $\rho = 0.4, 0.6, 0.9, 0.99$, and apply the model with ρ fixed at 0.99, allowing us to quantify the impact on model performance of misspecifying ρ. Additionally, we fit concordant models for each value of ρ to provide context for the misspecified settings.

To systematically introduce error into the dissimilarity metric, we generate misspecified values from a normal distribution centered at the true values, $\mathrm{N}(z_{ij}, \kappa)$, with $\kappa = 0, 1, 5, 10$. The generated dissimilarity metrics are constrained to [0-360°). We also fit concordant models based on the true dissimilarity metric values for context. The results from the simulation study are displayed in Table 1, where each reported estimate is based on 1,000 simulated datasets.

In Table 1, it is clear that estimation of the posterior mean of the CV of α_{tGH} (i.e., "ST CV") is robust when ρ is misspecified. This is consistent for all values of ρ in Table 1, as the EC remains near the nominal value of 0.95. Furthermore, the similar MSE values in the misspecified and concordant columns indicate that fixing ρ at 0.99 does not result in any loss in efficiency, though the values separate as the amount of misspecification increases. The same patterns are seen when error is added to the dissimilarity metric. These results indicate that the model is relatively robust to the misspecification of multiple inputs.

Table 1: The effect of misspecifying the spatial correlation parameter, ρ, and dissimilarity metric, z_{ij}, on model performance. The dissimilarity metric is generated from $\mathrm{N}(z_{ij}, \kappa)$ and constrained to [0-360). The table displays mean squared error (MSE) and empirical coverage (EC) for both misspecified and concordant settings. Each reported estimate is based on 1,000 simulated datasets.

	Misspecified		Concordant			Misspecified		Concordant	
True ρ	MSE	EC	MSE	EC	κ	MSE	EC	MSE	EC
0.99	–	–	0.060	0.98	0	–	–	0.039	0.97
0.90	0.056	0.94	0.052	0.98	1	0.042	0.99	0.042	0.98
0.60	0.085	0.96	0.074	0.99	5	0.047	0.96	0.044	0.98
0.40	0.115	0.95	0.095	0.99	10	0.053	0.95	0.043	0.97

4 Sensitivity Analysis in the VF Data Analysis

In Section 5 of the main text, we applied the introduced methodology to VF data. We assess sensitivity of those results to four modeling assumptions: i) the regularization upper bound for $\log(\alpha_{tGH})$, υ_1, ii) the temporal correlation structure, $\boldsymbol{\Sigma}(\phi)$, iii) the mean hyperparameter for $[\boldsymbol{\delta}]_1$, $[\boldsymbol{\mu}_\delta]_1$, and iv) the criterion for the bounds of ϕ. The "ST CV" metric is re-calculated under each sensitivity setting and regressed against the clinical assessment of progression with the results displayed in Table 2

Table 2: Sensitivity analysis to various assumptions in analysis of visual field data and glaucoma progression risk. For each assumption new values are presented with risk presented for comparison. Assumptions include i) upper bound for $\log(\alpha_{tGH})$, υ_1, ii) correlation structure, $\boldsymbol{\Sigma}(\phi)$, iii) mean hyperparameter for $[\boldsymbol{\delta}]_1$, $[\boldsymbol{\mu}_\delta]_1$, and iv) criterion for the bounds of ϕ.

Assumption	Value	Estimate	Std. Error	z value	Pr(>|z|)
υ_1	0.25	1.70	0.78	2.17	0.030
	0.50	2.03	0.84	2.43	0.015
	1.00	2.08	0.85	2.44	0.015
	2.00	2.02	0.86	2.35	0.019
$\boldsymbol{\Sigma}(\phi)$	EXP	2.08	0.85	2.44	0.015
	AR(1)	3.05	1.02	2.98	0.003
$[\boldsymbol{\mu}_\delta]_1$	0	2.07	0.85	2.44	0.015
	3	2.08	0.85	2.44	0.015
ϕ	0.99	2.07	0.82	2.52	0.012
	0.95	2.08	0.85	2.44	0.015

In the manuscript, the regularization upper bound for $\log(\alpha_{tGH})$ is set at 1.00, while

values of 0.25, 0.50, and 2.00 are used for comparison in the sensitivity analysis. The results are clearly robust to the choice of υ_1, as the p-values do not change. To demonstrate sensitivity to the exponential temporal correlation structure, we compare it to a continuous time autoregressive with lag one structure, AR(1), $[\boldsymbol{\Sigma}(\phi)]_{t,t'} = \phi^{|x_t - x_{t'}|}$. When using the AR(1) temporal correlation structure, it can be seen that "ST CV" is highly predictive of glaucoma progression (p-value: 0.0029), indicating that the results are robust over correlation structures. Next, we assess sensitivity to the specification of choosing $[\boldsymbol{\mu}_\delta]_1 = 3$, which was clinically motivated, comparing it to the objective choice of zero. The results are robust to the choice of $[\boldsymbol{\mu}_\delta]_1$, which is reassuring, but not surprising since the corresponding variance was 1,000. Finally, we compare the original results to a version where the lower bound of ϕ was based on the following criterion, $[a_\phi : [\boldsymbol{\Sigma}(a_\phi)]_{t,t'} = 0.99, |t - t'| = x_{\max}]$, changed from the original 0.95. This criterion is motivated to give wider and symmetric bounds. From Table 2, it is clear that the results are robust to changing this criterion. Overall, our model specification is robust to the modeling assumptions and prior choices in the VF data setting.

5 Additional Figure

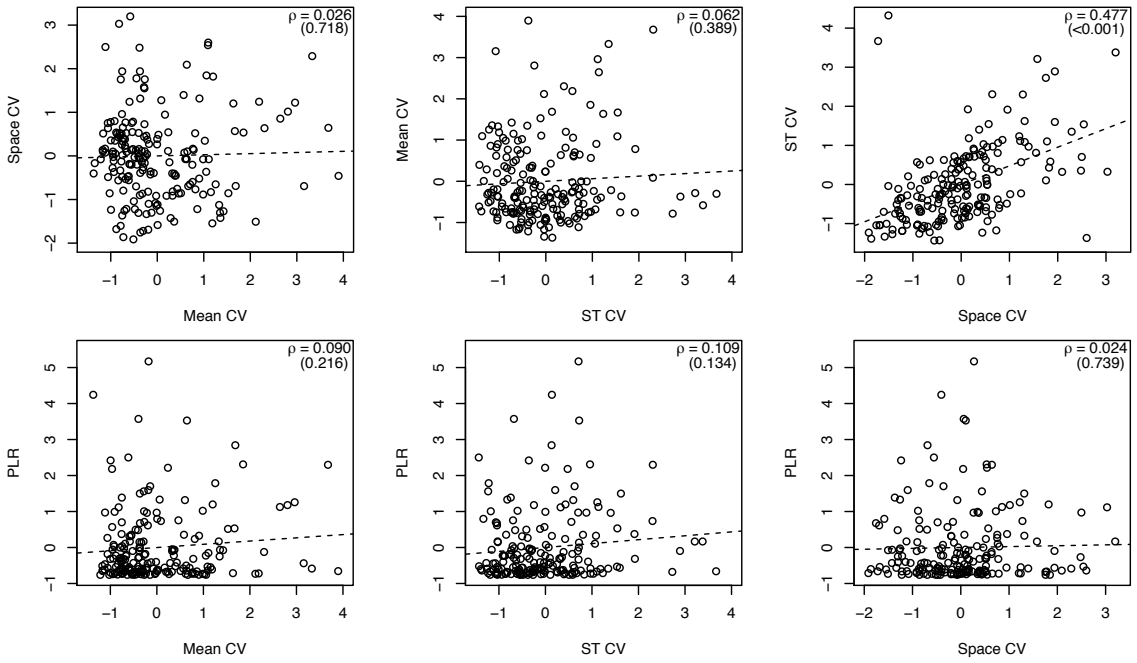

Figure 1: Pairwise scatterplots between standardized diagnostic metrics. Included are estimated Pearson correlations, corresponding p-values from the hypothesis that the correlations are equal to zero, and regression trend lines.

References

Banerjee, S., A. E. Gelfand, and B. P. Carlin (2003). *Hierarchical Modeling and Analysis for Spatial Data*. CRC Press.

Gelfand, A. E. and A. F. Smith (1990). Sampling-based approaches to calculating marginal densities. *Journal of the American Statistical Association 85*(410), 398–409.

Geman, S. and D. Geman (1984). Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence* (6), 721–741.

Geweke, J. (1992). Evaluating the accuracy of sampling-based approaches to calculating posterior moments. In J. M. Bernardo, J. Berger, A. P. Dawid, and J. F. M. Smith (Eds.), *Bayesian Statistics 4*, pp. 169–193. Oxford: Oxford University Press.

Metropolis, N., A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller (1953). Equation of state calculations by fast computing machines. *The Journal of Chemical Physics 21*(6), 1087–1092.