Bayesian Statistics (supplementary information)

Bayesian statistics differ from frequentist methods in that the unknown model parameters are given prior distributions before any data are observed. These distributions represent our complete prior knowledge of each parameter. Once data are observed, we update our prior beliefs by calculating the posterior distribution. Through this process, we are able to learn about the model parameters by combining our prior beliefs with the information contained in the data. When the prior distributions of the model parameters are very vague, containing little prior information, the results from the Bayesian analysis will closely match those of the corresponding frequentist analysis since a majority of the resulting information will be due to the observed data. A Bayesian analysis is preferred in this setting due to the efficient computational methods available to fit the proposed non-standard statistical change point model and the ability to properly characterize the uncertainty in the parameter estimates through the use of vague priors. To conduct inference on the unknown model parameters, the posterior distribution is summarized through use the posterior mean and 95% credible interval, providing a point estimate and measure of uncertainty for each parameter.

Bayesian Change Point Model 1 (CPM1):

The proposed model for the RNFL thickness for person i from a given sector is given as

$R_i = s_0 10^{0.1D_i} + b + \epsilon_i$ for $\theta \leq D_i \leq 0$; $R_i = s_0 + b + \epsilon_i$ for $D_i \geq 0$; and $R_i = s_0 10^{0.1\theta} + b + \epsilon_i$ for $D_i \leq \theta$.

The RNFL thickness for person i is given as R_i and the sensitivity loss for person i is given as D_i. The slope is given as s_0, the intercept term is given as b, and the person specific variability is given as ϵ_i where the ε_i have independent and identically distributed normal distributions such that $\epsilon_i \sim N(0, \sigma^2)$. We assume that for $D_i \geq 0$, the RNFL thickness is constant. This represents normal total deviation values and as a result there should be no change in the RNFL thickness. The RNFL then decreases with deteriorating retinal sensitivity loss through the equation $s_0 10^{0.1D_i} + b$. Finally, after a certain

value of retinal sensitivity loss, the RNFL thickness no longer decreases and becomes constant, reaching the end of the thinning process. This point of retinal sensitivity loss is given by θ in the proposed model, also known as the change point in the regression model. The thickness when RNFL stops thinning is then given by $s_0 10^{0.1\theta} + b$. Through the introduction of θ we are able to statistically estimate this change point. This model is repeated separately for each included sector and the global average.

To complete the change point model specification, we assign prior distributions to the unknown model parameters. s_0 and b are given independent normal distributions centered at zero with a large variance ($s_0, b \sim \text{Normal}(0, 100)$). This results in a rather uninformative prior distribution, essentially allowing the data alone to determine the value of these parameters. The variance parameter, σ^2, is given a vague inverse gamma prior distribution resulting in conjugacy in the model ($\sigma^2 \sim \text{Inverse Gamma}(0.01, 0.01)$). This prior is also relatively uninformative and places the emphasis on the data. Finally, the change point, θ, is given a uniform prior over the possible range of values. This range of values includes zero and the smallest observed value of total deviation ($\theta \sim \text{Uniform}(-35, 0)$). We assume that the RNFL thickness becomes flat at some value between those points. Using a uniform prior distribution allows each value in the range to be equally likely values for θ before the data are observed. This model is fit separately for each sector of interest and globally.

Bayesian Change Point Model 2 (CPM2):

The proposed model for the RNFL thickness for person i from a given sector is given as

$R_i = s_0(D_i - \theta) + b + \epsilon_i$ for $\theta \leq D_i$; and $R_i = b + \epsilon_i$ for $D_i \leq \theta$ where each of the terms have been previously defined. The same prior distributions are also used in this analysis. This model allows for two lines with different slopes to be connected at the change point location for D_i. When D_i is smaller than the change

point, the line is flat (zero slope), indicating no additional thinning of the RNFL thickness. For values of D_i larger than the change point, the slope (s_0) is estimated by the data.

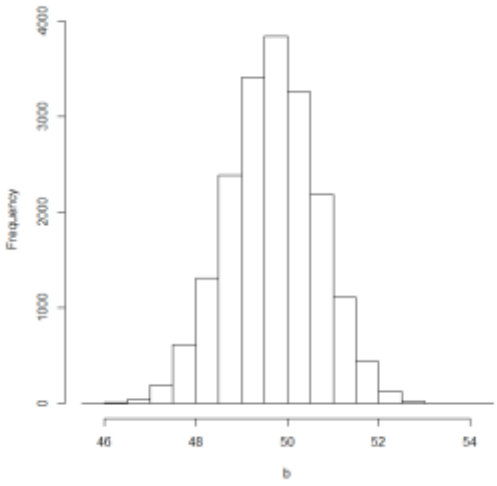
Markov chain Monte Carlo Details:

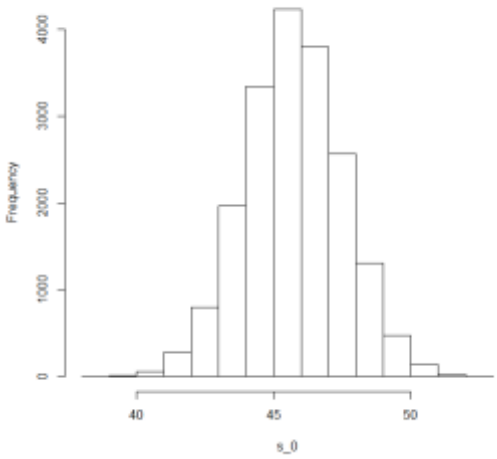
In order to fit the model, we rely on Markov chain Monte Carlo (MCMC) techniques. These MCMC sampling methods allow us to draw sample from the full posterior of interest for our model parameters. This full posterior distribution is given as

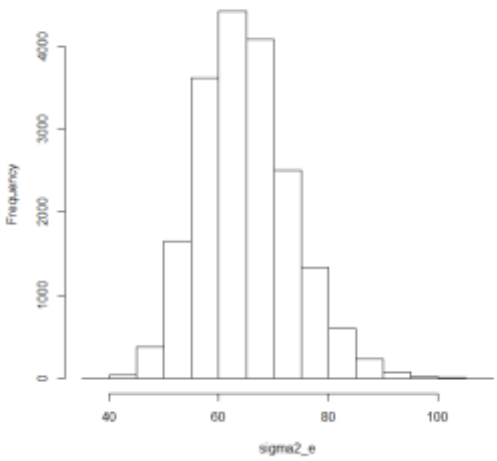
$$f(s_0, b, \sigma^2, \theta|\boldsymbol{R}) = \frac{f(\boldsymbol{R}|s_0, b, \sigma^2, \theta)f(s_0)f(b)f(\sigma^2)f(\theta)}{\int\int\int\int f(\boldsymbol{R}|s_0, b, \sigma^2, \theta)f(s_0)f(b)f(\sigma^2)f(\theta)ds_0 db \mathrm{d}\sigma^2 d\theta}$$

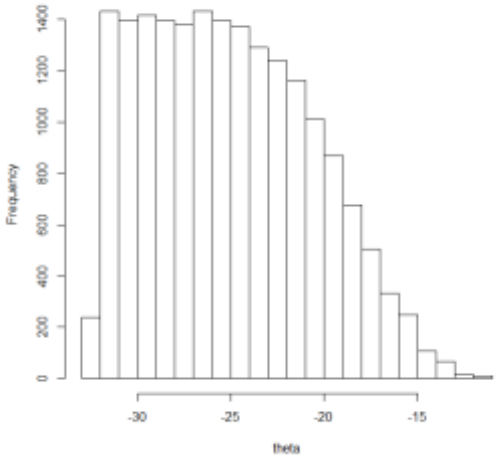
where $\boldsymbol{R}$ represents the full vector RNFL thicknesses from all patients. The full conditional distributions of $s_0, b,$ and σ^2 have known forms that allow for the use of Gibbs sampling while Metropolis sampling is required for the update of θ. We collect 190,000 posterior samples after discarding the first 10,000 draws during the burn in period before convergence. We then summarize the posterior distributions in order to obtain estimates and credible intervals for the parameters. For point estimates, we rely on the posterior mean and median and calculate the posterior standard deviation in order to present a measure of uncertainty for each parameter.

For the global Spectralis CPM1 fit, we display histograms of the posterior distribution for each model parameter. These are the samples that are summarized in order to make proper inference on the parameters. Other model fits similarly resulted in samples from the posterior distribution of each model parameter.


Frequency
0
1000
2000
3000
4000
46
48
50
52
54
b


Frequency
0
1000
2000
3000
4000
40
45
50
s_0


Frequency
0
1000
2000
3000
4000
40
60
80
100
sigma2_e


Frequency
0
200
400
600
800
1000
1200
1400
-30
-25
-20
-15
theta