Bayesian Hierarchical Modelling of Longitudinal Glaucomatous Visual Fields Incorporating Influential Factors

S.R. Bryan[1,2], K. A. Vermeer[2], P.H.C. Eilers[1], H.G. Lemij[3], E.M.E.H. Lesaffre[1,4]

1. Department of Biostatistics, Erasmus MC, Rotterdam, The Netherlands
2. Rotterdam Ophthalmic Institute, Rotterdam, The Netherlands
3. Glaucoma Service, Rotterdam Eye Hospital, Rotterdam, The Netherlands
4. KU Leuven, L-Biostat, Leuven, Belgium

Aim: Evaluation of a longitudinal series of visual fields (VF) provides a method to quantify functional deterioration in glaucoma patients. Sensitivity estimates are measured at 52 test locations within each eye (excluding the blind spot), and describe the level of differential light sensitivity, measured from 0 dB (blind) to 40 dB (upper end of normal). Factors such as test reliability, time of day and season have been shown to influence measurements. In addition, the horizontal raphe divides the retina in its upper and lower hemispheres. Consequently, correlation of VF measurements between hemispheres is lower than within hemispheres. Furthermore, the data is censored at 0 dB by the limited sensitivity range of the instrument. By investigating these aspects, we aim to develop a model which incorporates the high measurement variability, and can be used to improve predictions of point-wise VF data in clinical practice.

Methods: The motivating dataset consisted of VFs of 139 glaucoma patients from the Rotterdam Eye Hospital, measured approximately every 6 months over an average period of 9.6 years. One eye per individual was included in the analysis. A Bayesian hierarchical mixed effects model (M1) was used, including individual specific as well as location specific effects. As a further step, test reliability, time of day and season were included as covariates (M2). This was then extended to incorporate a spatial aspect by allowing the distributions for the locations belonging to the upper and lower retina to differ, and including censoring at 0 dB (M3). The models were compared using the deviance information criterion (DIC).

Results: M2, including test reliability, time of day and season, had a better fit (DIC=734907) compared to M1, which excludes these factors (DIC=735205). M3, including the spatial aspect and censoring, performed the best (DIC=734884). Using the latter model, the population averaged mean (and CI) was 19.75 dB (18.70; 20.67) and variance was 12.33 dB (12.24; 12.43), with an average slope of -0.16 dB/year (-0.32; -0.02). The sensitivity estimates increased by 0.04 dB (0.02; 0.05) per fixation loss, and 0.15 dB (0.12; 0.18) per false positive answer. Sensitivity estimates were 0.29 dB lower (-0.37; -0.21) when measured between noon and 2pm and 0.38 dB lower (-0.46; -0.31) when measured after 2pm, compared to those measured in the early morning. Measurements taken in autumn were 0.06 dB (0.01; 0.12) higher than those taken in spring.

Conclusion: We proposed a method to model point-wise VFs taking into account the hierarchical structure of the data. Test reliability, time of day and season were all shown to be significant factors that should be included. Extending this model to include the spatial aspect and censoring showed a further improvement in the model fit. Hence, this model is advantageous in dealing with the high measurement variability, and could be extended for the prediction of future VFs.