# Global Visit Effects in Point-Wise Longitudinal Modeling of Glaucomatous Visual Fields

Susan R. Bryan,[1,2] Paul H. C. Eilers,[2] Emmanuel M. E. H. Lesaffre,[2,3] Hans G. Lemij,[4] and Koenraad A. Vermeer[1]

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

**PURPOSE.** One of the difficulties in modeling visual field (VF) data is the sometimes large and correlated measurement errors in the point-wise sensitivity estimates. As these errors affect all locations of the same VF, we propose to model them as global visit effects (GVE). We evaluate this model and show the effect it has on progression estimation and prediction.

**METHODS.** Visual field series (24-2 Full Threshold; 15 biannual VFs per patient) of 125 patients with primary glaucoma were included in the analysis. The contribution of the GVE was evaluated by comparing the fitting and predictive ability of a conventional model, which does not contain GVE, to such a model that incorporates the GVE. Moreover, the GVE's effect on the estimated slopes was evaluated by determining the absolute difference between the slopes of the models. Finally, the magnitude of the GVE was compared with that of other measurement errors.

**RESULTS.** The GVE model showed a significant improvement in both the model fit and predictive ability over the conventional model, especially when the number of VFs in a series is limited. The average absolute difference in slopes between the models was 0.13 dB/y. Lastly, the magnitude of the GVE was more than three times larger than the measureable factors combined.

**CONCLUSIONS.** By incorporating the GVE in the longitudinal modeling of VF data, better estimates may be obtained of the rate of progression as well as of predicted future sensitivities.

Keywords: global visit effects, sensitivity estimates, model fitting, prediction, progression

Evaluation of a longitudinal series of visual fields (VF), as measured by standard automated perimetry (SAP), provides a method to detect glaucoma, and to determine functional deterioration.[1] One of the difficulties in modeling VF data is the large measurement variability of VFs, partially due to the inherent subjective nature of such a test.[2-4] This large variability means that in clinical practice, repeated measurements are performed to confirm real progression.

A learning effect, where the average mean defect significantly decreases between exams, has been shown to be present in VF testing.[2,5-7] Furthermore, test–retest studies have shown that variability is dependent on defect depth and test location.[2]

In glaucoma, variability is presumably related to fatigue effects and response errors. A fatigue effect, whereby sensitivity estimates decrease during an examination, has been demonstrated. This effect has been shown to be different between the inferior and superior hemifields within one eye.[8] In addition, this effect may differ between the first and second eye at the same visit. The number of false-negative answers have been shown to be higher in eyes with field loss.[9] It has also been shown that there is an inverse relationship between variability and sensitivity.[10] That is, there is a large amount of variability in eyes with severe damage.

A common approach to reduce measurement variability is to average multiple measurements. For example, summary measures such as the mean deviation (MD) have less variability due to the averaging over the point-wise sensitivity estimates. Other errors, however, are spatially correlated and affect the whole VF. One group of such errors are measurable factors, including season, time of day and reliability indices, which have been evaluated before.[11,12] Although these factors are statistically significant, they are rather small, and hence only explain a small part of the observed global variation in VFs. Junoy Montolio et al.[11] modelled the visit effect with these known factors. However, we speculate that other transient factors, such as fatigue, lack of concentration, or delayed reaction time may play a more important role. An example of the importance of these factors can be seen in Figure 1, where all locations have a drastic decrease in sensitivity in one of a series of visits. From the longitudinal profiles, it is evident that this decrease is caused by something that affected all VF measurements of that visit, rather than by actual damage.

Because all these (as well as possibly other) factors, affect all locations of the same VF, we propose to take them together and to call and model them as global visit effects (GVE). In this way, we can account for both the known and the unknown factors. Hence, the GVE accounts for all factors that affect all
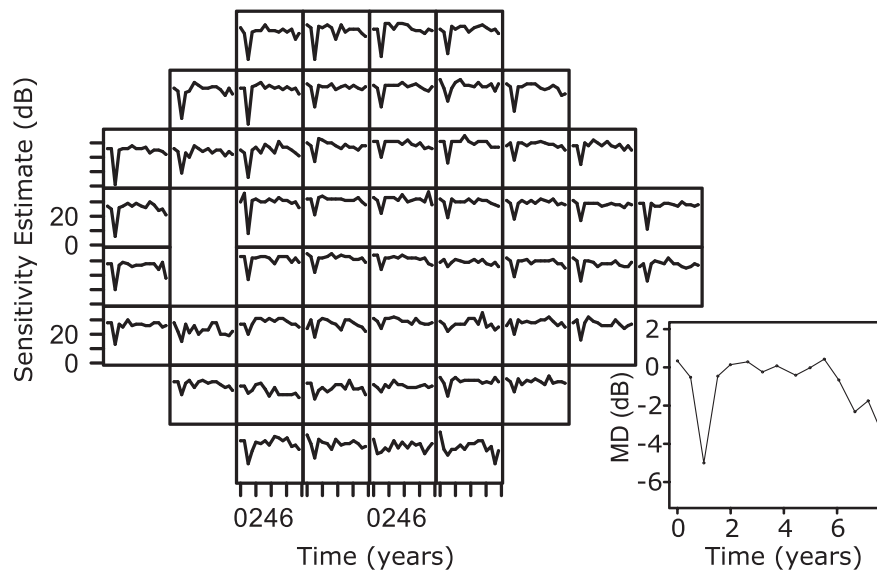
**FIGURE 1.** Retinal sensitivity estimates over time for each location of the visual field in the left eye of a single glaucoma patient. A decrease in the sensitivity estimates can be seen in all locations at around 1 year. The longitudinal profile of the MD values over time are shown on the right. The visit-dependent decrease is also clear at around 1 year for the MD.

measurements of the same eye at each visit. The GVE was inspired by a similar phenomenon, namely, trend estimation of sea levels for large numbers of monitoring stations.

We evaluated the GVE by determining the improvement of the model fit due to incorporating the GVE in the VF model and the effect of including the GVE on estimating the rate of progression. Furthermore, we determined the improvement of point-wise predictions for future measurements accounting for the GVE. Finally, we investigated the magnitude and importance of the GVE by comparing it with influential factors that have been discussed in the literature.

## METHODS

### Patients and Data

The analysis was performed on a subset of VF data of individuals from an ongoing study conducted at the Rotterdam Eye Hospital, The Netherlands; a full description of the data was provided earlier.[13,14] All data is available through the Rotterdam Ophthalmic Data Repository at http://rod-rep.com. In brief, the patients were followed up approximately twice per year. The VFs were tested by using the Humphrey Field Analyzer (Carl Zeiss Meditec, Dublin, CA, USA) with the 24-2, white-on-white test strategy by means of the Full Threshold algorithm. The response variables of interest were the sensitivity estimates from the 52 VF points (excluding the 2 points that correspond to the blind spot). All patients gave their written informed consent for participation. The research procedures followed the tenets set forth in the Declaration of Helsinki. We excluded VFs with unknown reliability as indicated by the instrument. Additionally, to simplify the evaluation of the statistical models, we excluded individuals with less than 15 measurements (in either eye). For those individuals with more than 15 measurements, only the first 15 measurements were included in the analysis. The resulting data set consisted of 250 eyes from 125 individuals, resulting in 3750 VFs and 195,000 location-specific sensitivity estimates. Descriptive statistics can be found in Table 1.

## Statistical Modeling

Bayesian models have many advantages over frequentist approaches, such as taking into account the uncertainty in all parameters. Combined with Markov chain Monte Carlo computations, they also allow for greater flexibility by relaxing the strong parametric assumptions commonly used in frequentist models. A hierarchical model is able to take into account both the within subject and between subject variability. Moreover, because both eyes were included in the analysis, we are able to capitalize on the common features within each eye by taking into account the correlation between measurements belonging to the same eye. In addition, any correlation of VF measurements within the inferior and superior hemifields, separated by the horizontal meridian, was expected to be higher than between hemifields.[15] Hence, the analysis was done by using a Bayesian hierarchical mixed-effects model.[16-18] We modeled the hierarchical structure of the data using four levels, namely, (1) the individual, (2) the eye, (3) the hemisphere, and (4) the location. An example of the mixed-effects model for the four level data structure can be seen in Figure 2A. Furthermore, censoring was taken into account at 0 dB,[19] due to the limitation of the device.[20] We will refer to this model as the conventional model. To account for the visit-dependent offset at all locations, or GVE, we included a parameter in the model to capture the offset at every visit for each eye within each individual. Hence, this effect accounts for factors that affect all measurements belonging to the same eye at each visit. The impact of this additional parameter is demonstrated in Figure 2B. This model will be referred to as the GVE model.

**TABLE 1.** Descriptive Statistics of the Study Sample

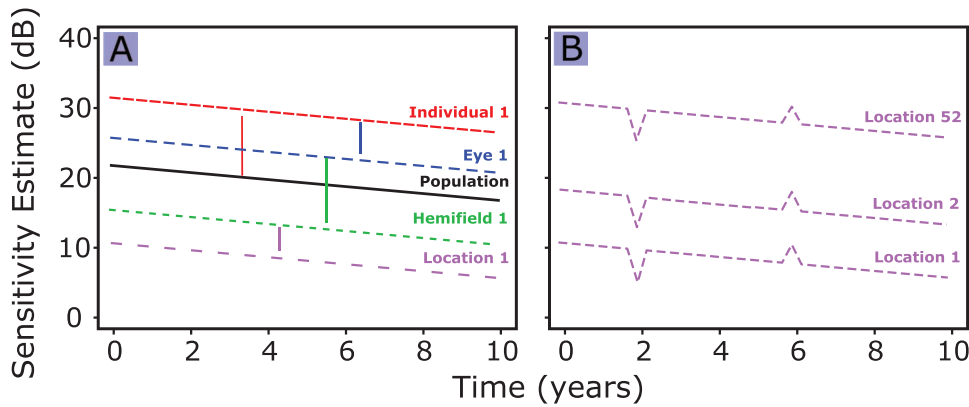|  | Mean | Median | Interquartile Range |
|---|---|---|---|
| Baseline age, y | 59.7 | 61.3 | 53.2; 67.1 |
| Baseline MD, dB | −7.8 | −5.7 | −11.7; −2.3 |
| Average change in MD, dB/y | −0.07 | −0.04 | −0.19; 0.13 |
| Follow-up time, y | 7.7 | 7.6 | 7.2; 8.1 |

**FIGURE 2.** (**A**) The lines represent the estimated evolutions for each of the levels using a hierarchical mixed effects model. In this example, only the intercept varies, but the model includes slopes in a similar way. (**B**) The GVE works as an offset for all locations belonging to the same VF. This offset is visible at around 2 years and 6 years in this example.

In classical, one-stage Bayesian hierarchical models, all parameters are estimated simultaneously. For complex models, obtaining results using this approach can be difficult or even impossible due to computational issues. In our case, problems were encountered regarding the running time and computer memory limitations. Thus, the analysis was done by using a recently proposed two-stage approach,[21] which allowed us to simplify the computation by splitting the hierarchical model at the individual level. Hence, individuals were analyzed independently before combining them at the population level. Figure 3 illustrates the hierarchical structure divided into the two stages. A full description of the models and the computational procedure is given by Bryan et al.[22]

### Model Evaluation

Our aim was to investigate different aspects of the GVE. Namely, how including the GVE affects the model fit, the estimated progression rate, the prediction of future measurements and the magnitude and importance of the GVE. To do this, we compared the conventional model with the GVE model. The models were compared by determining the error for each sensitivity estimate (predicted minus observed) and by combining these errors into summary error measures, namely, the root mean squared error (RMSE) and mean absolute error (MAE) for each model. The 95th percentile of the absolute errors, which is the value below which 95% of the absolute prediction errors may be found, was also computed to compare the models. Nonparametric Wilcoxon (matched paired when applicable) tests were performed to determine whether the differences between the models were significant.

### Model Fit

We evaluated the contribution of the GVE by comparing the model fits, for the conventional model and the GVE model, using the RMSE, the MAE and the 95th percentile of the absolute errors. An example of the fits for one eye can be seen in Figure 4.

### Effect on Estimated Progression Rate

To compare the estimated progression rate, or slope, of the conventional model and the GVE model, we first needed to correct for the GVE in the latter model. An example of this is shown in Figure 5. The evaluation of the change in slopes was

then done by calculating the mean absolute difference of the slopes between the conventional and the GVE model. Additionally, the distribution as well as the 95% confidence interval (CI) of the difference between the slopes of the two models was determined.

### Prediction of Future Visual Fields

To assess the predictive ability of the model, we performed a 5-fold cross validation. In each fold, the training set included 100 individuals while the testing set contained 25 individuals. Future measurements were then predicted for each location in the VF for each individual in the testing set. The number of measurements used for the estimation of the individual-specific effects was varied (3, 6, and 9 measurements), each time predicting the sensitivity estimate six measurements (~3 years) ahead. An example of this can be seen in Figure 6. In each of the examples, the GVE model gave a better prediction for the future measurement. The GVE model also appeared to be more robust over the number of measurements included. The models were evaluated by using the RMSE and the 95th percentile of the absolute errors.
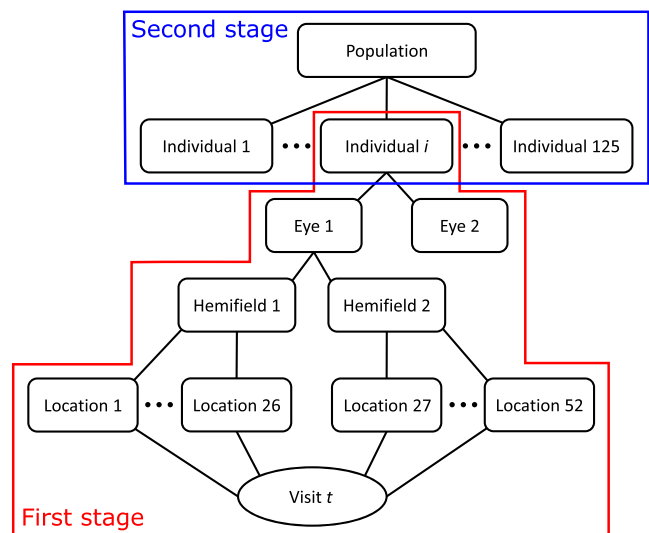


**FIGURE 3.** Illustration of the hierarchical structure of the data divided into the first and second stages as done in the two-stage approach.
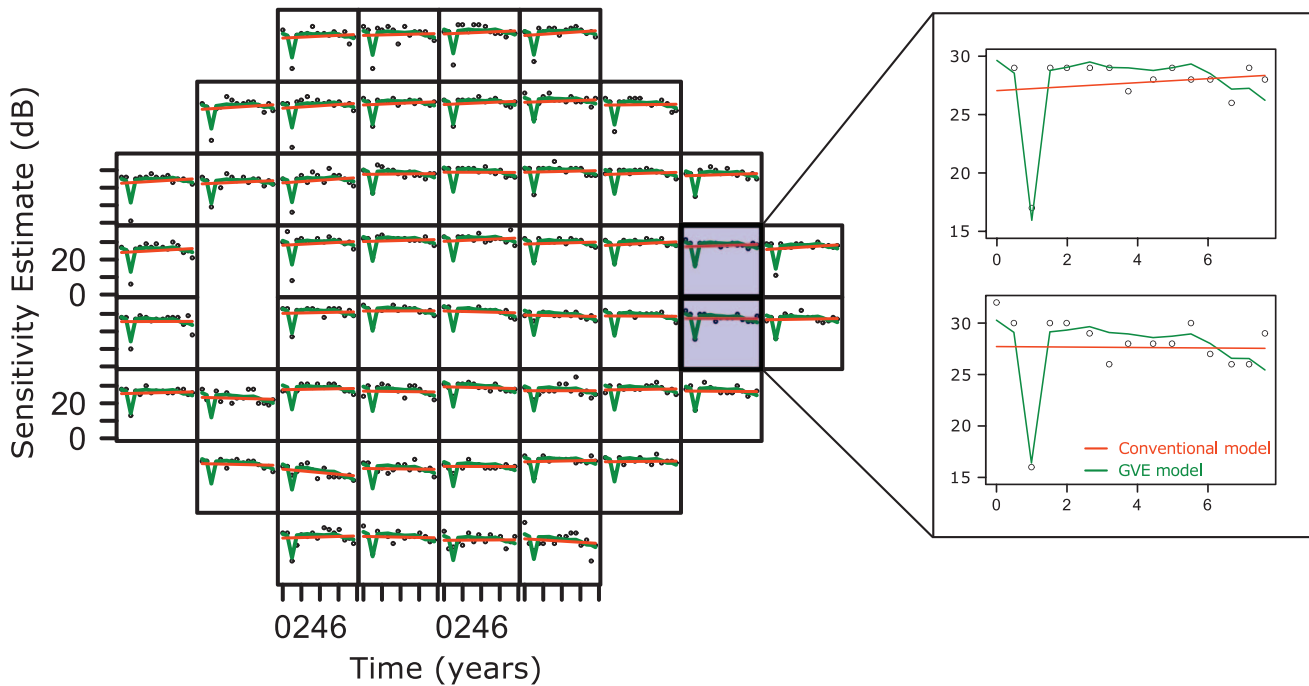
**FIGURE 4.** Scatter plot representing the retinal sensitivity estimates over time for each location of the VF in an example (left) eye. The lines represent the model fits for the conventional model (*orange*) and GVE model (*green*).

## Magnitude of the GVE

Junoy Montolio[11] showed that the time of day, season, reliability indices (number of fixation losses, false negatives, and false positive), technical experience, and follow-up period have a clinically relevant influence on the MD test results. To determine the magnitude and importance of the GVE, we compared it with these factors. We excluded technical experience and follow-up period due to the lack of data on these factors. Hence, we focussed on the time of day, season, and the reliability indices. For time of day, the tests were stratified into four categories: performed before 10 AM, between 10 AM and 12 PM, between 12 PM and 2 PM, and after 2 PM. For season, the tests were also stratified into four categories, of 3 months each (March–May, June–August, September–November, December–February), based on the annual variation of retinal sensitivity.[11] The reliability

indices were treated as continuous variables. Reliability indices include the percentage of fixation losses, the proportion of false positives, as well as the proportion of false negatives. We will refer to this model as the fixed-factor model. An example of the model fits for one location can be seen in Figure 7. We compared the model fits using the RMSE, MAE, and 95th percentile of the absolute errors. Furthermore, we determined the magnitude of the GVE compared with the factors by calculating their absolute means. A limitation of the two-stage approach occurs when there is sparse data, such as the season or time of day. Because each individual was analyzed separately, information could not be borrowed from the data set as a whole as done in the classical one-stage approach. Due to this limitation, we used the classical one-stage approach including 50 randomly selected individuals for this analysis.
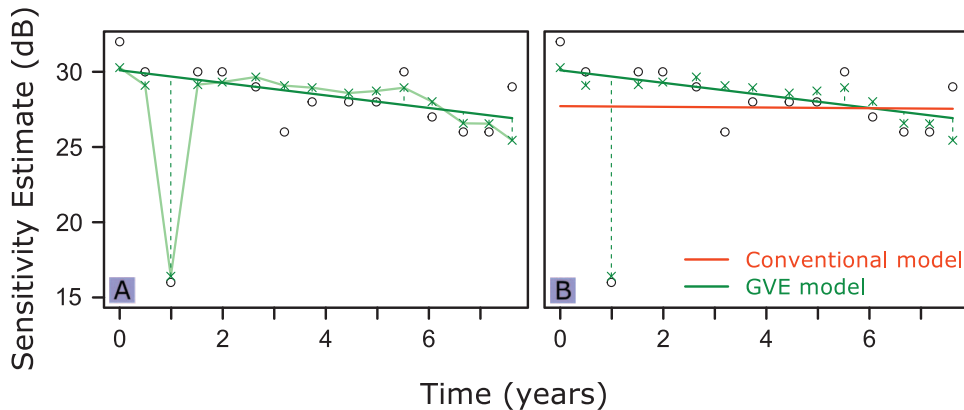


**FIGURE 5.** A real example showing how the GVE can influence the estimated rate of progression. The open circles represent the retinal sensitivity estimates over time for 1 location in the VF. (**A**) The crosses represent the fitted data using the GVE model, with the *transparent green line* showing the model fit over time. The *solid green line* indicated the slope of the GVE model. (**B**) The *orange line* shows the fit for the conventional model. The *green line* shows the fit for the GVE model after correcting for the GVE estimates.
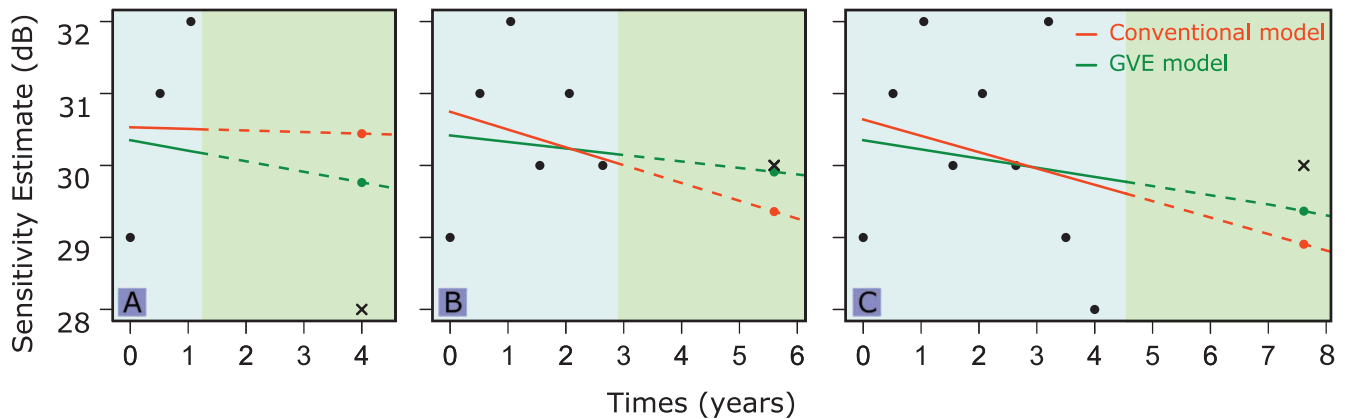
**FIGURE 6.** Scatterplots represents the retinal sensitivity estimates over time for 1 location of the VF. The predictions were done by using (**A**) 3, (**B**) 6, and (**C**) 9 measurements, which are shown as points in the *blue portion* of each plot. In each example, the sensitivity estimate 6 measurements ahead, represented by a cross, was predicted. The lines represent the predictions using 3, 6, and 9 measurements to predict 6 measurements ahead for the conventional model (*orange*) and GVE model (*green*).

## RESULTS

### Model Fit

Table 2 lists the RMSEs, MAEs, and 95th percentile of the absolute errors for the models, showing that by incorporating the GVE there is an improvement in the model fit. Both the squared errors and absolute errors were significantly smaller for the GVE model than for the conventional model ($P < 0.001$, matched paired).

### Effect on Slopes

When comparing the rate of progression between the conventional model and the GVE model, the mean absolute difference was found to be 0.13 dB/y ($P < 0.001$, matched paired). Figure 8 shows the distribution of the differences, including the mean and 95% CI.

### Prediction of Future Visual Fields

Table 3 lists the RMSE by using 3, 6, and 9 measurements to predict VF sensitivities approximately 3 years ahead (9th, 12th, and 15th measurement, respectively). The GVE model showed a significant improvement in the predictions compared with the conventional model, irrespective of how many measure-

ments were used ($P < 0.001$, matched paired). However, the difference between the models predictive abilities decreased as more measurements were included. For the conventional model, there was a significant difference between including three and six measurements ($P < 0.001$) and between including six and nine measurements ($P < 0.001$). For the GVE model, these differences were not significant ($P = 0.08$ and $P = 0.47$, respectively).

### Magnitude of the GVE

The fits for the conventional and the fixed-factor model were very similar. This can be seen in Table 4, which lists the RMSE, MAE, and 95th percentile of the absolute errors for each model. By including the factors, the model fit was slightly improved compared with the conventional model ($P < 0.001$, matched paired). The improvement in the fit was much larger for the GVE model compared with both the conventional ($P < 0.001$, matched paired) and the fixed-factor model ($P < 0.001$, matched paired).

The mean absolute value of each of the factors is shown in Table 5, with the most influential factor being season (0.13 dB). The combined value of all factors was found to be 0.23 dB. In contrast, the GVE was found to be 0.85 dB.

## DISCUSSION

In this study, we proposed to model measurement errors that affect the point-wise sensitivity estimates within the same VF as GVEs. By correcting for the GVEs, we accounted for all measureable factors, both known and unknown, that affect all measurements of the same eye at each visit. Furthermore, we evaluated this model and showed the magnitude of the correction for the GVE on progression estimation and prediction of future measurements.



**FIGURE 7.** Scatterplots representing the retinal sensitivity estimates over time for 1 location of the VF. The lines correspond to the fits of the 3 models, with much larger effects for the GVE (*green*) than for known, fixed factors (*blue*).

**TABLE 2.** Comparison of the Fitting Ability of the Models Using the RMSE, MAE, and the 95th Percentile of the Absolute Error

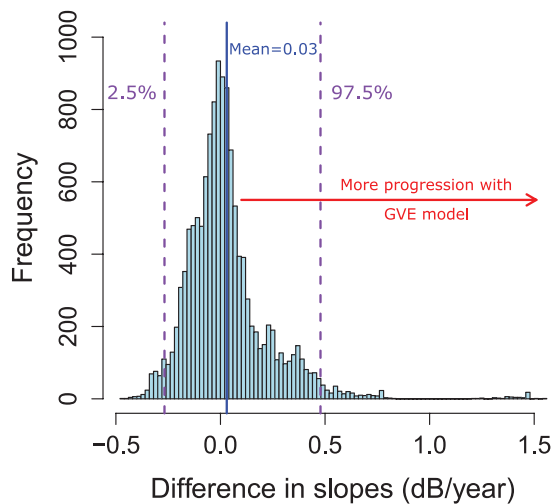| Model | RMSE, dB | MAE, dB | 95th Percentile, dB |
|---|---|---|---|
| Conventional model | 3.23 | 2.08 | 7.01 |
| GVE model | 2.97 | 1.91 | 6.43 |

**FIGURE 8.** Histogram showing the distribution of the differences between the slopes for the conventional model and the slopes for the GVE model. Positive values represent more progression (or smaller slopes) for the GVE model.

The GVE model showed a highly significant improvement in the model fit compared with the conventional model. Hence, by taking the GVE into account, we were able to remove a large systematic component that would otherwise be interpreted as noise and obtain better estimates of the true rate of progression. The mean absolute difference between the rates of progression estimated by the conventional model and the GVE model was shown to be 0.13 dB/y. The distribution was slightly positively skewed. By correcting for the GVE with our model, we found a predominantly higher rate of progression. In other words, ignoring the GVE may obscure the true, most often higher, rate of progression. In this analysis, all measurements were included to determine the difference between the progression rates of the two models.

For this model to be of clinical use, its prediction of future measurements should be better than of the conventional model. By incorporating the GVE in the model, we indeed found significant improvements in the predictions. The GVE model provided more robust predictions than the conventional model, especially when fewer measurements were included. This could be explained by the fact that the GVE model takes factors into account that may either decrease or increase the sensitivity estimates. The GVE has less effect on the estimated slopes when more measurements are included. Thus, the conventional model and the GVE model become more similar as more measurements are included. In clinical practice, the number of VFs is typically quite small, especially when resources are limited and robust approaches, such as the GVE, are therefore needed to accurately predict glaucomatous progression.

The importance of the GVE was compared with factors described in the literature that have been shown to be significant. The magnitude of the GVE, however, was shown

**TABLE 3.** Comparison of the Predictive Ability of the Models Using the RMSE, Varying the Number of Measurements Included for the Estimation

| Model | 3 Measurements, dB | 6 Measurements, dB | 9 Measurements, dB |
|---|---|---|---|
| Conventional model | 5.63 | 5.31 | 4.92 |
| GVE model | 4.58 | 4.65 | 4.67 |

**TABLE 4.** Comparison of the Fitting Ability of the Models Using the RMSE, MAE, and the 95th Percentile of the Absolute Error

| Model | RMSE, dB | MAE, dB | 95th Percentile, dB |
|---|---|---|---|
| Conventional model | 3.33 | 2.14 | 7.21 |
| Fixed-factor model | 3.32 | 2.14 | 7.18 |
| GVE model | 3.07 | 1.97 | 6.67 |

to be more than three times larger than that of the other factors combined (0.85 vs. 0.23 dB). Our results on those factors were consistent with those found by Junoy Montolio et al.[11] Of the known factors, we found season to have the largest effect, with a mean absolute value of 0.13 dB. In agreement with our results, Junoy Montolio et al.[11] concluded that the number of false positive answers has the largest, or most severe, effect out of the reliability indices. In their study, the MD was determined to be overestimated by 1 dB per 10% of false positive answers. Time of day was found to have a mean absolute value of 0.09 dB. It has, however, been shown that the 24-hour IOP rhythm differs between eyes in glaucoma patients.[23] Because the patient's IOP may affect each eye differently at the same time of day, and will vary between individuals, it may be of more interest to determine this factor at an eye-specific level, as done with the GVE. Although the GVE takes both measureable and unmeasurable factors into account, including the known factors allows for some explanation of the variation. Hence, a combination of both the GVE and the known factors may be beneficial in the modeling of VF data.

In our model, we determined the GVE for each VF. Thus, the GVE differs per eye for each individual at every visit. In this way, we can take factors into account that affect the two eyes (of the same individual) differently, such as, for example, the fatigue effect. Because the patient has VF tests done on both eyes sequentially, the second eye could be more affected by the lengthy test. Our model can be extended to take the correlation of the GVE between eyes into account. In that way, we could determine the patient-dependent, as well as the eye-dependent GVE. Furthermore, the model may be extended to accommodate that some locations may be affected differently by some of the factors. Although we have shown the effect of the GVE using an advanced Bayesian two-stage model, it is important to note that the GVE can be implemented as a random effect in any point-wise longitudinal model.

In comparison to conventional robust regression which is applied at the location level, the GVE model can be considered a robust approach applied at the VF level. Hence, while the GVE model is able to distinguish between within- and between-field errors, the conventional robust model is not able to do so, because both errors are combined in a single error term. By modelling the visit effect explicitly, we exploit the entire VF structure. Hence, the GVE allows us to obtain important

**TABLE 5.** Comparison of the Magnitudes of the Different Factors and the GVE Using the MAE

| Factor | MAE, dB |
|---|---|
| Season | 0.13 |
| Time of day | 0.09 |
| % Fixation losses | 0.00 |
| % False positives | 0.10 |
| % False negatives | 0.02 |
| Combined | 0.23 |
| Global visit effect | 0.85 |

additional information, which is needed in order to make any conclusions on the instrument, operator, and so on. In this way, we also allow for 'real' progression to still affect the slope of the model rather than attributing the entire measurement to error as done with the conventional robust model.

An approach that is related to the GVE is the pattern deviation, which uses the cumulative distribution to compute a correction.[20] The 85th percentile is used to compensate for effects by for instance cataract, which would lead to a general reduction of retinal sensitivity throughout the VF. Hence, the entire VF height is adjusted to the 85th percentile. However, with diffuse loss, the entire VF height tends to be overcorrected whenever the 85th percentile becomes significantly affected.[24] The two approaches differ in how the visit effects are expressed. Namely, the percentile correction is treated as a fixed term, while the GVE is treated as a random effect. Because we treat the GVE as a random effect, which has a distribution with mean zero, it is forced to fluctuate around zero. The fixed term of the percentile model does not impose any constraint, and hence accounts for both the visit effect and the rate of progression/slope. Hence, in contrast to pattern deviation analysis, the GVE model allows us to estimate the visit effect without disrupting the estimation of the progression.

In clinical practice it is known that one single VF may not be reliable and confirmation is sought in the results from future tests. Clustered VF testing has been shown to identify more rapid progression than evenly spaced follow-up approaches, allowing the clinician to be more confident that the progression is real.[25] Our model is in line with these findings and confirms that results, which deviate from what is expected may be due to unknown factors that affect the VF measurements on that specific visit rather than representing actual damage.

In conclusion, the GVE has a substantial effect on the pointwise VF sensitivities. In longitudinal VF series, correcting for the GVE provided better estimates of the true rates of progression, which were predominantly higher than with conventional progression analysis. In addition, shorter VF series were required to arrive at relatively accurate VF predictions than with conventional modeling. Implementing progression models that incorporate the GVE in clinical care may therefore improve the clinical management of glaucoma.

## Acknowledgments

## References

1. Flammer J, Meier E. *Glaucoma: A Guide for Patients: An Introduction for Care-Providers: A Quick Reference.* 2nd ed. Cambridge, MA: Hogrefe & Huber Publishers; 2003.

2. Heijl A, Lindgren G, Olsson J. The effect of perimetric experience in normal subjects. *Arch Ophthalmol.* 1989;107: 81–86.

3. Werner EB, Petrig B, Krupin T, Bishop KI. Variability of automated visual fields in clinically stable glaucoma patients. *Invest Ophthalmol Vis Sci.* 1989;30:1083–1089.

4. Kutzko KE, Brito CF. Effect of instructions on conventional automated perimetry. *Invest Ophthalmol Vis Sci.* 2000;41: 2006–2013.

5. Gloor B, Schmied U. Findings in glaucomatous visual fields under observation using the 'Octupus' automated perimeter. *Klin Monatsbl Augenheilk.* 1980;176:545–546.

6. Gloor B, Schmied U, Fassier A. Glaucomatous visual fields analysis of Octopus observation with statistical material. *Klin Monatsbl Augenheilk.* 1980;117:423–436.

7. Kulze JC, Stewart WC, Sutherland SE. Factors associated with a learning effect in glaucoma patients using automated perimetry. *Arch Ophthalmol.* 1990;68:681–686.

8. Hudson C, Wild JM, O'Neill EC. Fatigue effects during a single session of automated static threshold perimetry. *Invest Ophthalmol Vis Sci.* 1994;35:268–280.

9. Bengtsson B, Heijl A. False-negative responses in glaucoma perimetry: indicators of patient performance or test reliability? *Invest Ophthalmol Vis Sci.* 2000;41:2201–2204.

10. Russell RA, Crabb DP, Malik R, Garway-Heath DF. The relationship between variability and sensitivity in large-scale longitudinal visual field data. *Invest Ophthalmol Vis Sci.* 2012; 53:5985–5990.

11. Junoy Montolio FG, Wesselink C, Gordijn M, Jansonius NM. Factors that influence standard automated perimetry test results in glaucoma: test reliability, technician experience, time of day, and season. *Invest Ophthalmol Vis Sci.* 2012;53: 7010–7017.

12. Gardiner SK, Demirel S, Gordon MO, Kass MA; for The Ocular Hypertension Treatment Study Group. Seasonal changes in visual fields sensitivity and intraocular pressure in the Ocular Hypertension Treatment Study. *Ophthalmology.* 2013;120: 724–730.

13. Bryan SR, Vermeer KA, Eilers PH, Lemij HG, Lesaffre EM. Robust and censored modeling and prediction of progression in glaucomatous visual fields. *Invest Ophthalmol Vis Sci.* 2013; 54:6694–6700.

14. Erler NS, Bryan SR, Eilers PH, Lesaffre EM, Lemij HG, Vermeer KA. Optimizing structure-function relationship by maximizing correspondence between glaucomatous visual fields and mathematical retinal nerve fiber models. *Invest Ophthalmol Vis Sci.* 2014;55:2350–2357.

15. Asman P, Heijl A. Glaucoma hemifield test: automated visual field evaluation. *Arch Ophthalmol.* 1992;110:812–819.

16. Verbeke G, Molenberghs G. *Linear Mixed Models for Longitudinal Data.* 1st ed. New York: Springer; 2000.

17. Ntzoufras I. *Bayesian Modeling Using WinBUGS.* 1st ed. New York: Wiley & Sons; 2009.

18. Lesaffre E, Lawson A. *Bayesian Biostatistics.* 1st ed. New York: Wiley & Sons; 2012.

19. Tobin J. Estimation of relationships for limited dependent variables. *Econometrica.* 1958;26:24–36.

20. Anderson DR, Patella VM. *Automated Static Perimetry.* 2nd ed. St. Louis, MO: Mosby;1999.

21. Lunn D, Barrett J, Sweeting M, Thompson S. Fully Bayesian hierarchical modelling in two stages, with application to meta-analysis. *J R Stat Soc Ser C Appl Stat.* 2013;62:551–572.

22. Bryan SR, Eilers PH, Li B, et al. Bayesian hierarchical modeling of longitudinal glaucomatous visual fields using a two-stage approach. *ArXiv:*1502.03979[stat.AP].

23. Liu JHK, Weinreb RN. Asymmetry of habitual 24-hour intraocular pressure rhythm in glaucoma patients. *Invest Ophthalmol Vis Sci.* 2014;55:7398–7402.

24. Artes PH, Nicolela MT, LeBlanc RP, Chauhan BC. Visual field progression in glaucoma: total versus pattern deviation analyses. *Invest Ophthalmol Vis Sci.* 2005;46:4600–4606.

25. Crabb DP, Garway-Heath DF. Intervals between visual field tests when monitoring the glaucomatous patient: wait-and-see approach. *Invest Ophthalmol Vis Sci.* 2012;53:2770–2776.

Investigative Ophthalmology & Visual Science