The Usefulness of Assessing Glaucoma Progression With Postprocessed Visual Field Data

Sampson L. Abu1,2, Shervonne Poleon3, and Lyne Racette1

1 Department of Ophthalmology and Visual Sciences, University of Alabama at Birmingham, Birmingham, AL, USA
2 Pennsylvania College of Optometry, Salus University, Elkins Park, PA, USA
3 School of Optometry, University of Alabama at Birmingham, Birmingham, AL, USA

Correspondence: Lyne Racette, Department of Ophthalmology and Visual Sciences, University of Alabama at Birmingham, 1720 University Boulevard, Suite 415, Birmingham, AL 35294, USA. e-mail: lracette@uabmc.edu

Received: May 7, 2021
Accepted: April 18, 2022
Published: May 6, 2022

Keywords: glaucoma; postprocessed; progression; visual field

Citation: Abu SL, Poleon S, Racette L. The usefulness of assessing glaucoma progression with postprocessed visual field data. Transl Vis Sci Technol. 2022;11(5):5, https://doi.org/10.1167/tvst.11.5.5

Purpose: Data postprocessing with statistical techniques that are less sensitive to noise can be used to reduce variability in visual field (VF) series. We evaluated the detection of glaucoma progression with postprocessed VF data generated with the dynamic structure–function (DSF) model and MM-estimation robust regression (MRR).

Method: The study included 118 glaucoma eyes with at least 15 visits selected from the Rotterdam dataset. The DSF and MRR models were each applied to observed mean deviation (MD) values from the first three visits (V1–3) to predict the MD at V4. MD at V5 was predicted with data from V1–4 and so on until the MD at V9 was predicted, creating two additional datasets: DSF-predicted and MRR-predicted. Simple linear regression was performed to assess progression at the ninth visit. Sensitivity was evaluated by adjusting for false-positive rates estimated from patients with stable glaucoma and by using longer follow-up series (12th and 15th visits) as a surrogate for progression.

Results: For specificities of 80% to 100%, the DSF-predicted dataset had greater sensitivity than the observed and MRR-predicted dataset when positive rates were normalized with corresponding false-positive estimates. The DSF-predicted and observed datasets had similar sensitivity when the surrogate reference standard was applied.

Conclusions: Without compromising specificity, the use of DSF-predicted measurements to identify progression resulted in a better or similar sensitivity compared to using existing VF data.

Translational Relevance: The DSF model could be applied to postprocess existing visual field data, which could then be evaluated to identify patients at risk of progression.

Introduction

Glaucoma remains a leading cause of blindness. Although the disease progresses slowly in most patients, it progresses rapidly in others and results in clinically significant visual disability over a short period of time.1,2 Therefore, monitoring patients for early detection of glaucoma progression is the next crucial step following diagnosis and the initiation of treatment. Visual field (VF) assessment with static automated perimetry has been the clinical standard for monitoring functional changes in glaucoma.3,4 However, the results of VF tests, even when performed over short time intervals, can vary significantly. This presents a major challenge in monitoring glaucoma progression, because variability in VF data can either obscure or mimic deteriorating visual function.4,5 As a result, more frequent testing and longer follow-ups are needed to reliably assess progression.3,6

Different strategies for reducing VF variability have been explored. These include modifying conventional white-on-white automated perimetry by testing with a larger stimulus,7,8 using previous test results as thresholding priors,9 and the use of structure-guided testing.10–13 Postprocessing existing VF data has also been proposed as an alternative for minimizing the impact of VF variability on progression assessment. On the premise that VF sensitivities below 15 to 20 decibels (dB) are unreliable and not clinically useful for monitoring change,14 Gardiner et al.15 and Wall et al.16 investigated whether censoring unreliable sensitivities
would be detrimental to the detection of progression. Both studies found that censoring unreliable sensitivities did not hamper the ability to detect progression. Additionally, the application of statistical techniques such as filtering and spatial correlations to existing VF data has been shown to predict future VF status more accurately compared to using unprocessed VF measurements.\textsuperscript{17–23} One advantage of postprocessing available VF data is that no additional testing is required.

Ordinary least squares linear regression (OLSLR) has been the conventional method for evaluating longitudinal VF data.\textsuperscript{24,25} However, the prediction accuracy of OLSLR and its ability to identify progression can be severely compromised by outliers,\textsuperscript{4,26} which may result from learning effect, rim artifact, technician error, patient inattentiveness during testing, and/or increased test–retest variability associated with advanced disease. Robust regression, on the other hand, is less vulnerable to outliers\textsuperscript{18,26,27} and could provide a more reliable evaluation of longitudinal VF data. In Taketani et al.,\textsuperscript{18} future VF sensitivities were predicted more accurately by robust regression compared to OLSLR. The dynamic structure–function (DSF) model is another approach that has demonstrated better prediction accuracy than OLSLR for short data series.\textsuperscript{28,29} Both robust regression and the DSF model can be applied to existing VF data to generate series of predicted VF measurements. By design, the series of measurements generated with these predictive models would have lower variability compared to the original VF series. Therefore, their use could provide a more reliable assessment of progression and present the opportunity for early detection. In the current study, we evaluated the performance of assessing glaucoma progression with postprocessed measurements obtained with the DSF model and robust regression.

**Methods**

**Datasets**

Two datasets were analyzed in the current study. The first dataset was drawn from the Rotterdam Ophthalmic Data Repository which is publicly accessible at [http://www.rodrep.com/data-sets.html](http://www.rodrep.com/data-sets.html). This repository contains de-identified longitudinal VF data for 139 glaucoma patients recruited from the Rotterdam Eye Hospital in the Netherlands.\textsuperscript{19} Each patient had glaucomatous visual loss defined as the presence of any two of the following: (1) abnormal pattern standard deviation value, abnormal glaucoma hemifield test result, and a cluster of at least three points depressed at the $P = 0.05$ level, or (2) at least one point depressed at the $P = 0.01$ level. A detailed description of the eligibility criteria for the Rotterdam Eye Study can be found elsewhere.\textsuperscript{19} We also analyzed a de-identified test–retest VF dataset for patients with stable glaucoma and well-controlled levels of intraocular pressure. These patients were recruited from the glaucoma clinics at the Queen Elizabeth Health Sciences Centre in Halifax, Nova Scotia.\textsuperscript{30} This dataset is publicly accessible as part of the R visual Fields package.\textsuperscript{31} Both studies were conducted in accordance with the tenets of the Declaration of Helsinki. Research protocols were approved by the respective institutional review boards, and all patients provided written informed consent.

**Visual Field Testing**

For the Rotterdam Eye Study, patients were tested every 6 months using the 24-2 full threshold program of the Humphrey Visual Field Analyzer (Carl Zeiss Meditec, Dublin, CA). The study eye of each patient in the test–retest dataset was tested weekly over a period of 3 months (totaling 12 VF sessions) using the 24-2 SITA Standard program of the Humphrey Field Analyzer.

**Inclusion Criteria for the Current Study**

From the Rotterdam dataset, we selected an eye from each patient with at least 15 visits. Patients with baseline VF mean deviation (MD) worse than –15 dB were excluded. When both eyes of a patient were eligible, one eye was randomly chosen. For the test–retest dataset, we included only patients who completed all 12 VF tests within a 3-month period. Baseline information for each dataset is summarized in the Table.

**Generation of Postprocessed Datasets With the DSF Model and Robust Regression**

Two separate predicted datasets were generated by applying the DSF model and robust regression to the observed MD values from the Rotterdam dataset. Figure 1 illustrates the range of visits from which the predicted MD series were generated for each model. Although the DSF model was originally developed to use structural and functional information jointly,\textsuperscript{28,29} prediction of only VF measurements is feasible by keeping the structural input constant. The model combines two vectors—a centroid and velocity vector—to estimate future measurements. The centroid
describes the state of the disease and is computed as the average of observations available at a given time point. The velocity vector estimates the direction and average rate of change in measurements over time. In the current study, the DSF model was applied to observed MD values from the first three visits (V1–3), measured at times $t_1$, $t_2$, and $t_3$, to predict the MD value for V4 at time $t_4$. The model first determined the centroid ($C$) as $\frac{MD_1 + MD_2 + MD_3}{3}$, and then calculated the velocity vector ($V$) as $\frac{MD_1 - MD_2}{t_1 - t_2} + \frac{MD_2 - MD_3}{t_2 - t_3}$. The MD value at V4 was estimated as the sum of the centroid and the expected change in measurements between $t_3$ and $t_4$, given as $C + V(t_4 - t_3)$. This procedure was repeated to predict MD values for V5 through V9 using the available observed MD values from the respective preceding visits: V1–4, V1–5, V1–6, V1–7, and V1–8. This set of six predicted MD values is subsequently referred to as the DSF-predicted data.

Next, we generated another set of six predicted MD values using the MM-estimation robust regression (MRR). Deriving its name from two underlying maximum likelihood methods (the M- and S-estimations), the MRR technique combines a high breakdown point and high statistical efficiency. This technique involves the weighting of each data point based on the magnitude of its residual, where larger residuals are assigned smaller weights. This is an iterative procedure that seeks to identify outliers and minimize their impact on the coefficient estimates. First, MRR was performed on observed MD values from V1–3, and the expected MD value at V4 was extrapolated from the line of best fit. We repeated this procedure on observed MD values from V1–4, V1–5, V1–6, V1–7, and V1–8 to estimate MD values at V5, V6, V7, V8, and V9, respectively. This set of predicted MD values generated with robust regression is subsequently referred to as the MRR-predicted data.

Table. Baseline Description of the Two Datasets Included in This Study

| Dataset                        | Rotterdam Dataset | Test–Retest Dataset |
|--------------------------------|-------------------|--------------------|
| Eligible patients, $n$         | 118               | 29                 |
| Age (y), mean (SD)             | 58.4 (10.2)       | 66.3 (8.2)         |
| Follow-up (mo), mean (SD)      | 92.4 (7.2)        | 2.9 (0.4)          |
| Baseline MD (dB), mean (SD)    | -3.9 (4.2)        | -3.1 (3.2)         |

Figure 1. Illustration of data prediction with the DSF model and robust regression. For each model, we used MD values from the first three visits (V1–3) to predict MD at V4, and then MD values from V1–4, V1–5, V1–6, V1–7, and V1–8 to predict MD values at V5, V6, V7, V8, and V9, respectively.

Statistical Analysis

Using simple linear regression, we assessed progression at the ninth visit with the DSF-predicted dataset, the MRR-predicted dataset, and the available observed measurements (MD values from V1–V9). Figure 2 shows five case examples of progression assessment with each dataset to exemplify how each predicted dataset compared with the observed dataset. Due to the non-existence of a gold standard for glaucoma progression, assessing sensitivity is challenging and can only be done through indirect measures. One approach is to determine the false-positive estimates associated with a given method from a cohort of stable glaucoma eyes and then use these estimates to adjust the positive or hit rates (the proportion of eyes flagged as progressing). Alternatively, a surrogate reference standard can be established from the evaluation of longer VF series, against which performance of a new method can be assessed. In the current study, both approaches were
Figure 2. Case examples of progression evaluation using each dataset. (A) A case where progression was not flagged using any of the datasets. (B) A case where progression was identified with only the observed data. (C) Progression was identified with all datasets. (D) A case where progression at the ninth visit was flagged with only predicted data and confirmed at subsequent visits using the observed data. (E) Similar to panel D, except that progression was not confirmed at the 12th and 15th visits.
applied to determine the sensitivity of using predicted measurements to assess progression.

In the first approach, we used a test–retest dataset (stable glaucoma eyes) to obtain false-positive rates associated with the DSF-predicted and MRR-predicted datasets. For each eye, the series of observed MD values from the first nine tests was selected and shuffled 35 times to increase the sample size from 29 to 1015. Each shuffled set of observed data was then used to generate the DSF-predicted and MRR-predicted MD values in a similar fashion as previously described. The shuffling of the datasets disrupted the temporal sequence of the MD series, thereby eliminating any potential statistically significant negative slope in the measurements. Moreover, the test–retest dataset was obtained over a short time frame within which no clinically significant change was expected. Therefore, after evaluating the DSF-predicted, MRR-predicted, and observed MD values with simple linear regression, the proportion of stable eyes flagged as showing a statistically significant rate of change counted as their respective false-positive rates. Supplementary Figure S1 shows the false-positive rates obtained at 0 to 0.20 significance levels for the DSF and MRR models. Finally, the sensitivities of identifying progression with the DSF-predicted, MRR-predicted, and observed datasets were determined by normalizing the positive rates obtained for the evaluation of eyes in the Rotterdam dataset. The normalization was achieved by adjusting the positive rates by the corresponding false-positive rates derived from the test–retest dataset. For example, at a specificity of 95%, if the DSF-predicted dataset yielded a 15% positive rate in the Rotterdam dataset and a 5% false-positive rate in the test–retest dataset, then its sensitivity would be 10% at that given specificity (i.e., 15% positive rate minus 5% false positive rate).

As mentioned previously, we also examined the performance of each dataset in reference to a surrogate gold standard for progression. This reference standard was defined based on the statistical significance of the observed MD slopes at the 12th and 15th visits (i.e., longer follow-up series). Compared to nine visits, assessing progression over 12 or 15 visits offered a more reliable estimate of the status and rate of progression. Receiver operating characteristic (ROC) curves and the associated partial area under the curve (pAUC) were generated to compare the performance of predicted measurements to that of observed measurements. We acknowledge that this surrogate gold standard is imperfect and incapable of flagging some eyes that might show progression if longer follow-up length was available; for example, some eyes were flagged by both predicted measurements models but not by the observed measurements, as illustrated in Figure 2E. Such eyes would be counted as false positives and consequently lead to underestimation of the area under the curve for the predicted dataset. Therefore, we also estimated positive rates for each dataset at significance levels of 0 to 1 as an alternative measure of their ability to detect progression. The positive rate was computed as the proportion of eyes correctly identified as progressing based on the progression outcome at the 12th and 15th visits.

To determine the accuracy of the rate of progression estimated with the predicted datasets, we evaluated the difference between the observed MD slopes at the 15th visit and the predicted MD slopes for increasing length of follow-up (seventh to 15th visit). Furthermore, the prediction error on future VF measurements was obtained for the DSF model and MRR model. The prediction error for each model was computed by subtracting the predicted measurement from a more stable “observed value” estimated from linear regression of all observed MD values available at the 12th visit and 15th visit. The mean absolute error (MAE) for each prediction was compared between the models. Data prediction and all analyses were performed using R, with robust regression implemented by the lmrob function of the robustbase R package and the ROC curves generated by the pROC package.

Results

Figure 3 presents a comparison of the ability of the DSF-predicted dataset, MRR-predicted dataset, and the observed dataset to identify progression at the ninth visit. Of the three datasets, the DSF-predicted dataset had the greatest sensitivity at specificity of 80% to 100%. Out of the 118 eyes in the Rotterdam dataset, 27 eyes (22.9%) and 36 eyes (30.5%) showed a statistically significant change in observed MD values at the 12th visit and 15th visit, respectively. Figure 4 presents ROC curves depicting the sensitivities of the three datasets computed in reference to progression outcomes at the 12th and 15th visits (surrogate gold standard). The sensitivity of the DSF-predicted dataset was similar to that of the observed dataset, with pAUCs of 0.84 (95% confidence interval [CI], 0.75–0.93) for the observed dataset and 0.69 (95% CI, 0.59–0.81) for the DSF-predicted dataset when the 12th visit was used as a surrogate for progression. Similarly, when the 15th visit was used as a surrogate, the pAUCs were 0.68 (95% CI, 0.59–0.78) for the observed dataset and 0.60 (95% CI, 0.53–0.70) for the DSF-predicted dataset. The
Figure 3. Comparison of the performance of the DSF-predicted dataset, MRR-predicted dataset, and the observed dataset in detecting progression at the ninth visit. Sensitivity estimates for each dataset were normalized using the corresponding false-positive rates derived from the test–retest dataset (a cohort of stable glaucoma patients).

The MRR-predicted dataset was the least sensitive, with the lowest pAUCs and 95% CIs (0.50–0.66).

Figure 5 presents the accuracy of the rates of progression obtained using the DSF- and MRR-predicted datasets. The rates of progression obtained with these datasets at increasingly longer follow-up periods are compared with those obtained using the longest follow-up period (15 visits) available in the observed data, which provides the most accurate estimates. Even though the range of the difference in rate of progression computed for each predicted dataset became narrower with increasing series length (appearing to converge at the 15th visit), the average interquartile range (IQR) for the DSF-predicted dataset was smaller than for the MRR-predicted dataset (0.24 dB/y vs. 0.63 dB/y). The MAEs for the prediction of MD value at the 12th visit were 0.61 dB for the DSF model and 0.73 dB for the MRR model. These values were 0.77 dB for the DSF model and 1.20 dB for the MRR model for the prediction of MD value at the 15th visit.

Discussion

Monitoring functional change in glaucoma is challenging given the inherent variability in VF measurements. This variability can reach levels that violate the assumptions about measurement error underlying classical least squares regression, thereby yielding inaccurate evaluation of longitudinal VF data. By way of weighting each data point and by using centroids (average of available measurements), estimates from robust regression and the DSF model are respectively less influenced by measurement variability. This property was leveraged in the current study by utilizing each model to postprocess existing VF data. When the false-positive rates derived from patients with stable glaucoma were used to correct for the positive rate obtained in the Rotterdam dataset, the sensitivity of the DSF-predicted dataset was higher than that of both the observed and MRR-predicted dataset (Fig. 3). The sensitivity of the DSF-predicted dataset was, however, similar to that obtained with the observed dataset when longer follow-up series were used as a surrogate reference standard for progression (Fig. 4). These results suggest that the analysis of DSF postprocessed data could be useful in screening glaucoma patients to identify those likely to progress and estimate future progression status in individual patients. This could enhance our ability to detect glaucoma progression early and to better plan for clinical interventions ahead of time.

Although less than one-fifth of glaucoma patients under care would progress rapidly, undetected slow progression can be detrimental to a younger patient over time. A screening tool that would help identify patients at risk of rapid progression without missing slow progression would be useful in glaucoma management. Despite having similar classification accuracy (based on pAUC) as the observed dataset (Fig. 4), the ability of the DSF-predicted dataset to flag relatively more eyes as progressing (Fig. 3) at a fixed specificity can be harnessed to develop a screening tool for glaucoma progression. Although the higher sensitivity of this approach suggests the possibility overcalling of progression, a false alert that a patient may be at greater risk of progression would only result in closer monitoring, which would not be as detrimental as misdiagnosing a non-glaucomatous eye for treatment. The greater sensitivity associated with using the DSF-predicted dataset could be an inflation of the false-positive rate inherent in the data prediction process. However, this was ruled out with the evaluation of the test–retest dataset in which the false-positive estimates for each predicted dataset equaled the proportion of eyes expected to be flagged as progressing due to chance (Supplementary Fig. S1). With a median coefficient of determination ($R^2$) of 0.70 (IQR, 0.58) for the DSF-predicted dataset, 0.12 (IQR, 0.30) for the observed dataset, and 0.43 (IQR, 0.69) for MRR-predicted dataset, factors such as reduced variability within the DSF-predicted dataset potentially account for its greater sensitivity.
Figure 4. ROC curves showing the sensitivities and the associated pAUC (95% CI) for each dataset in reference to a surrogate gold standard for progression. Sensitivities were derived in reference to progression outcomes at the 12th visit (left panel) and 15th visit (right panel).

In the absence of a standardized reference for glaucoma progression, in one approach we used progression outcomes based on observed data available at the 12th and 15th visits as a surrogate gold standard. Among the limitations of this surrogate reference is the possibility that the series of 12 or 15 VFs may be insufficient follow-up data to confirm or rule out the possible progression flagged by predicted datasets in some eyes. For example, in Figure 2E progression was flagged by both the DSF-predicted (slope = −0.09 dB/y; \( P < 0.01 \)) and the MRR-predicted measurements (slope = −0.41 dB/y; \( P = 0.01 \)) but was not identified in the observed data available at the 15th visit (slope = −0.02 dB/y; \( P = 0.33 \)). These misclassifications between predicted and observed datasets may or may not hold when data from a longer period are available. Given the follow-up length in the study, we observed that both predicted datasets tended to flag more stable eyes as progressing compared to the observed dataset, as presented in Supplementary Figure S2A. This ultimately can affect the sensitivity and pAUC estimated for the predicted datasets. For example, the MRR-predicted dataset was the least sensitive approach (Fig. 4) despite having about 30% less within-series variability compared to the observed dataset. Given that data prediction did not compromise specificity (Supplementary Fig. S1), patients flagged as progressing with the evaluation of predicted measurements may be at greater risk of progression and could benefit from close monitoring, especially when longer follow-up data are not available.

Clinicians currently rely on repeated testing and longer follow-up to obtain reliable data for accurate assessment of progression. However, this increases the burden on patients and clinic resources and may further delay the detection of progression. In contrast, postprocessing existing VF data by means of predicting future measurements with statistical models offers a practical and inexpensive solution for reducing measurement variability without requiring additional testing. We observed that the DSF model predicted MD values more accurately than the MRR model, with MAE differences of 0.12 dB and 0.43 dB for predictions at the 12th and 15th visits, respectively. We further observed that the evaluation of DSF-predicted measurements yielded comparable rates of progression as obtained with longer observed VF series (Fig. 5). These findings suggest that evaluation of postprocessed VF data derived with the DSF model could provide a comparatively reliable estimate of the rate of progression early in some patients.

In contrast to approaches that are based on either population estimates such as glaucoma progression analysis or censoring unreliable sensitivities with predetermined cut-offs, analysis of postprocessed data may permit individualized progression assessment, as these predicted measurements are generated from each patient's own existing data. A commonly used model for individualized progression assessment—permutation of pointwise linear regression (PoPLR)—involves estimation of the overall significance of deterioration across all VF locations in reference to 5000 permutations of a patient's own data. The PoPLR model has been shown to be more sensitive to functional progression than trend analysis of a single global index such as MD. We performed PoPLR analysis on series of raw VF sensitivities from
Figure 5. Accuracy of the rates of progression estimated with predicted measurements. The box plots show the difference in rate of progression between the predicted datasets as generated using increasingly longer follow-up periods (x-axis) and the observed dataset at visit 15 (the longest follow-up period). Panels A and B show the differences in rate of progression obtained for the DSF-predicted dataset and the MRR-predicted dataset, respectively.

V₁ to V₉ extracted from the Rotterdam dataset and compared the proportion of eyes it flagged as progressing to that obtained with the analysis of predicted and observed MD values (Fig. 6). Consistent with the findings of O’Leary et al., the PoPLR model flagged a greater proportion of eyes as progressing compared to conventional trend analysis of observed MD values. Analyses of DSF-predicted and MRR-predicted global MD measurements had greater positive rates compared to the PoPLR model. The generation and analysis of pointwise postprocessed VF measurements represent an avenue of future work necessary to determine whether it could further improve sensitivity to functional progression.

As test–retest variability increases with disease severity, the magnitude of prediction error is expected to become larger. The centroid of the DSF model is an average estimation of the level of damage, and its relationship with prediction error could be useful for modeling the expected variability in measurements as the disease progresses. For each DSF prediction, we determined the associated prediction error as the absolute difference between the observed value and the predicted value. A simple linear regression of the absolute differences on the corresponding centroids revealed that 2% to 15% of the variability in predictor error was accounted for by the centroid. This suggests that the centroid may not be useful for modeling expected variability. Nonetheless, because the centroid is less sensitive to variability in observed measurements, it allows for the derivation of less variable predicted measurements, which could be clinically useful even in advanced disease.

In this study, the DSF model consistently included the first three data points (MD values from V₁–V₃) in making all predictions. This could potentially result in conservative predictions and shallower rates of progression in cases where extreme values occur earlier in the series. To examine the impact of this on our findings, we explored a “moving window” alternative,
where predicted measurements were obtained by applying the DSF model to only the set of three preceding data points. Thus, we used MD values from V₁ to V₃ to predict V₄, V₂ to V₄ to predict V₅, V₃ to V₅ to predict V₆, and so on. At a specificity of 95%, MD values predicted using either all available data or based on a “moving window” flagged 37 eyes, 34 of which were flagged by both methods, yielding a kappa agreement of 0.88 (substantial agreement). This finding suggests that including the first data points in all DSF predictions did not significantly alter the ability to identify change in this study. Predictions based on a “moving window” may be appropriate when extreme values occur early in the observed series, or in the clinical setting where the most recent values are most relevant.

The current study has limitations. First, the length of MD series used to assess progression at the ninth visit differed between the observed and predicted datasets. Our rationale for analyzing all nine observed MD values from V₁ to V₉ versus six predicted MD values for V₄ to V₉ was based on the fact that observed data for V₁–₃ were used in generating all predicted measurements. Although longer series provide a more accurate estimate of progression, the DSF-predicted dataset was more sensitive compared to the observed dataset at the ninth visit. Even fewer eyes were flagged as progressing when shorter series of observed data (V₁–₆ or V₄–₉) were analyzed. Another limitation is that the patients enrolled in the Rotterdam Eye Study were under standard clinical care, and those who showed signs of deterioration likely received modifications to their treatment to slow the rate of progression. Given that the observed and predicted datasets obtained within the same period were analyzed and compared, we do not believe that variability in the rate of progression over time introduced systematic bias in favor of either of the methods used in this study.

In conclusion, we have demonstrated in this study that assessing progression with postprocessed VF measurements generated with the DSF model resulted in similar or better sensitivity compared to observed data and that it yielded comparable rates of progression as longer observed VF series, without compromising specificity. These findings may be due to the reduced variability within the DSF-predicted series of measurements. In the absence of a widely accepted method to identify glaucoma progression, the evaluation of postprocessed measurements may be useful for identifying patients at greater risk of progression. These patients could be monitored more closely to determine whether changing or intensifying therapy is warranted to prevent or slow glaucomatous vision loss.

Acknowledgments

The authors are grateful to the Rotterdam Eye Institute for making the visual field data publicly available.

Supported by a grant from the National Institutes of Health (EY025756 to LR) and an unrestricted grant from Research to Prevent Blindness.

Disclosure: S.L. Abu, None; S. Poleon, None; L. Racette, Olleyes, Inc. (C)
References

1. Chauhan BC, Malik R, Shuba LM, Rafuse PE, Nicolela MT, Artes PH. Rates of glaucomatous visual field change in a large clinical population. Invest Ophthalmol Vis Sci. 2014;55(7):4135–4143.
2. Heijl A, Buchholz P, Norrgren G, Bengtsson B. Rates of visual field progression in clinical glaucoma care. Acta Ophthalmol. 2013;91(5):406–412.
3. Chauhan BC, Garway-Heath DF, Gohi FJ, et al. Practical recommendations for measuring rates of visual field change in glaucoma. Br J Ophthalmol. 2008;92(4):569.
4. Hu R, Racette L, Chen KS, Johnson CA. Functional assessment of glaucoma: uncovering progression. Surv Ophthalmol. 2020;65(6):639–661.
5. Artes PH, Nicolela MT, LeBlanc RP, Chauhan BC. Visual field progression in glaucoma: total versus pattern deviation analyses. Invest Ophthalmol Vis Sci. 2005;46(12):4600–4606.
6. Crabb DP, Russell RA, Malik R, et al. Frequency of Visual Field Testing When Monitoring Patients Newly Diagnosed with Glaucoma: Mixed Methods and Modelling. Southampton, UK: NIHR Journals Library; 2014.
7. Wall M, Brito CF, Woodward KR, Doyle CK, Johnson CA. Total deviation probability plots for stimulus size v perimetry: a comparison with size III stimuli. Arch Ophthalmol. 2008;126(4):473–479.
8. Gardiner SK, Demirel S, Goren D, Mansberger SL, Swanson WH. The effect of stimulus size on the reliable stimulus range of perimetry. Trans Vis Sci Technol. 2015;4(2):10.
9. Turpin A, Jankovic D, Funk RD, Johnson CA. How many visual fields are required to precisely predict future test results in glaucoma patients when using different trend analyses? Invest Ophthalmol Vis Sci. 2015;56(6):4076–4082.
10. Bryan SR, Vermeer KA, Eilers PHC, Lemij HG, Lesaffre EMEH. Robust and censored modeling and prediction of progression in glaucomatous visual fields. Invest Ophthalmol Vis Sci. 2013;54(10):6694–6700.
11. McNaught AI, Crabb DP, Fitzke FW, Hitchings RA. Modelling series of visual fields to detect progression in normal-tension glaucoma. Graefes Arch Exp Ophthalmol. 1995;233(12):750–755.
12. Betz-Stablein BD, Morgan WH, House PH, Hazelton ML. Spatial modeling of visual field data for assessing glaucoma progression. Invest Ophthalmol Vis Sci. 2013;54(2):1544–1553.
13. Gardiner SK, Crabb DP, Fitzke FW, Hitchings RA. Reducing noise in suspected glaucomatous visual fields by using a new spatial filter. Vision Res. 2004;44(8):839–848.
14. Spry PG, Johnson CA, Bates AB, Turpin A, Chauhan BC. Spatial and temporal processing of threshold data for detection of progressive glaucomatous visual field loss. Arch Ophthalmol. 2002;120(2):173–180.
15. Bengtsson B, Patella VM, Heijl A. Prediction of glaucomatous visual field loss by extrapolation of linear trends. Arch Ophthalmol. 2009;127(12):1610–1615.
16. Russell RA, Crabb DP, Malik R, Garway-Heath DF. The relationship between variability and of the macula using structural information. Transl Vis Sci Technol. 2018;7(6):36.
17. Wall M, Woodward KR, Doyle CK, Zamba G. The effective dynamic ranges of standard automated perimetry sizes III and V and motion and matrix perimetry. Arch Ophthalmol. 2010;128(5):570–576.
18. Morales E, de Leon JMS, Abdollahi N, Yu F, Nouri-Mahdavi K, Caprioli J. Enhancement of visual field predictions with pointwise exponential regression (PER) and pointwise linear regression (PLR). Trans Vis Sci Technol. 2016;5(2):12.
19. Taketani Y, Murata H, Fujino Y, Mayama C, Asaoka R. How many visual fields are required to precisely predict future test results in glaucoma patients when using different trend analyses? Invest Ophthalmol Vis Sci. 2015;56(6):4076–4082.
sensitivity in large-scale longitudinal visual field data. *Invest Ophthalmol Vis Sci.* 2012;53(10):5985–5990.

26. Wilcox RR. *Introduction to Robust Estimation and Hypothesis Testing.* 3rd ed. Amsterdam: Academic Press; 2012.

27. Huber PJ, Ronchetti EM. *Robust Statistics.* 2nd ed. Hoboken, NJ: John Wiley & Sons; 2009.

28. Hu R, Marin-Franch I, Racette L. Prediction accuracy of a novel dynamic structure-function model for glaucoma progression. *Invest Ophthalmol Vis Sci.* 2014;55(12):8086–8094.

29. Abu SL, KhalafAllah MT, Racette L. Evaluation of the external validity of a joint structure–function model for monitoring glaucoma progression. *Sci Rep.* 2020;10(1):19701.

30. Artes PH, O’Leary N, Nicolela MT, Chauhan BC, Crabb DP. Visual field progression in glaucoma: what is the specificity of the Guided Progression Analysis? *Ophthalmology.* 2014;121(10):2023–2027.

31. Marin-Franch I, Swanson WH. The visualFields package: a tool for analysis and visualization of visual fields. *J Vis.* 2013;13(4):10.

32. Gardiner SK, Crabb DP. Examination of different pointwise linear regression methods for determining visual field progression. *Invest Ophthalmol Vis Sci.* 2002;43(5):1400–1407.

33. Zhu H, Crabb DP, Ho T, Garway-Heath DF. More Accurate Modeling of Visual Field Progression in Glaucoma: ANSWERS. *Invest Ophthalmol Vis Sci.* 2015;56(10):6077–6083.

34. Montesano G, Garway-Heath DF, Ometto G, Crabb DP. Hierarchical censored Bayesian analysis of visual field progression. *Transl Vis Sci Technol.* 2021;10(12):4.

35. R Core Team. *R: A Language and Environment for Statistical Computing.* Vienna, Austria: R Foundation for Statistical Computing; 2020.

36. O’Leary N, Chauhan BC, Artes PH. Visual field progression in Glaucoma: estimating the overall significance of deterioration with permutation analyses of pointwise linear regression (PoPLR). *Invest Ophthalmol Vis Sci.* 2012;53(11):6776–6784.

37. Gardiner SK, Demirel S, De Moraes CG, et al. Series length used during trend analysis affects sensitivity to changes in progression rate in the ocular hypertension treatment study. *Invest Ophthalmol Vis Sci.* 2013;54(2):1252–1259.