Time Lag Between Functional Change and Loss of Retinal Nerve Fiber Layer in Glaucoma

Stuart K. Gardiner, Steven L. Mansberger, and Brad Fortune

Devers Eye Institute, Legacy Research Institute, Portland, Oregon, United States

Correspondence: Stuart K. Gardiner, Devers Eye Institute, Legacy Research Institute, 1225 NE 2nd Avenue, Portland, OR 97210, USA; sgardiner@deverseye.org.

Received: June 12, 2020
Accepted: October 11, 2020
Published: November 3, 2020

Citation: Gardiner SK, Mansberger SL, Fortune B. Time lag between functional change and loss of retinal nerve fiber layer in glaucoma. *Invest Ophthalmol Vis Sci.* 2020;61(13):5. https://doi.org/10.1167/iovs.61.13.5

Purpose. It is often suggested that structural change is detectable before functional change in glaucoma. However, this may be related to the lower variability and hence narrower normative limits of structural tests. In this study, we ask whether a time lag exists between the true rates of change in structure and function, regardless of clinical detectability of those changes.

Methods. Structural equation models were used to determine whether the rate of change in function (mean linearized total deviation, $AveTD_{Lin}$) or structure (retinal nerve fiber layer thickness [RNFLT]) was predicted by the concurrent or previous rate for the other modality, after adjusting for its own rate in the previous time interval. Rates were calculated over 1135 pairs of consecutive visits from 318 eyes of 164 participants in the Portland Progression Project, with mean 207 days between visits.

Results. The rate of change of $AveTD_{Lin}$ was predicted by its own rate in the previous time interval, but not by rates of RNFLT change in either the concurrent or previous time interval (both $P > 0.05$). Similarly, the rate of RNFLT change was not predicted by concurrent $AveTD_{Lin}$ change after adjusting for its own previous rate. However, the rate of $AveTD_{Lin}$ change in the previous time interval did significantly improve prediction of the current rate for RNFLT, with $P = 0.005$, suggesting a time lag of around six months between changes in $AveTD_{Lin}$ and RNFLT.

Conclusions. Although RNFL thinning may be detectable sooner, true functional change appears to predict and precede thinning of the RNFL in glaucoma.

Keywords: glaucoma, diagnostics, progression, structure-function relation, structural equation models

Several previous studies have suggested that structural changes may be detectable prior to functional changes in glaucoma.[1–4] However, the focus of those studies was not on which modality shows *true change* first, but on which shows *detectable change* first. This is complicated by the issue of variability, which is higher for standard automated perimetry (SAP) than for retinal nerve fiber layer thickness (RNFLT).[5] Indeed, the World Glaucoma Association's consensus document states the following:

> With current technology, detection of structural defects generally precedes detectable functional defects in glaucoma patients while functional defects can precede structural defects in some patients… Structural tests based on the comparison to the normative data tend to show a statistically significant glaucomatous change earlier compared to the functional tests because of a greater variability in functional tests.[6]

It is likely that the structure-function relation would be much tighter in the absence of this test-retest variability.[7] As technology develops and variability is reduced, the temporal relation between being able to detect structural and functional changes could be altered, despite the course of underlying pathophysiology following the same sequence.

At the cellular level, there are strong reasons to hypothesize that functional change may actually precede structural changes in human glaucoma. Dead or terminally nonfunctional neurons may still occupy physical space and thus contribute to measurements of structure but not function; whereas functioning neurons must occupy some volume and thus contribute to measurements of both structure and function. Indeed, in experimental glaucoma, axon loss of around 12% was found in the optic nerves of eyes without any reduction in RNFLT compared with their contralateral control eye.[8,9]

In this study, we ask whether there is evidence of a time lag between changes in structure (RNFLT in this case) and function (SAP) in human glaucoma. We aim to identify the temporal relation between *true changes* in these modalities, not just *detectable changes*; that is, without regard to whether those changes are outside the normal limits of test-retest variability. To achieve this, we use structural equation models (SEM), which allow us to examine the strength of the predictive relation between the two while accounting for unobservable latent variables, correlations between predictors, and autocorrelated repeated measures. By using biannual testing, this technique requires that any time lag will not

just be statistically significant, but several months in magnitude. Such a lag could then be considered important for our understanding of the disease process and also inform future developments to diagnostic testing.

Methods

Participants and Data Collection

Data were used from participants enrolled in the ongoing Portland Progression Project, a longitudinal study of progression and diagnostic testing in glaucoma.[10,11] Inclusion criteria were a diagnosis of open-angle glaucoma or likelihood of developing glaucoma, as judged by the participant's clinician, to reflect a typical clinical population. Exclusion criteria were a history of angle closure, presence of other ocular pathologies likely to affect the visual field (e.g., diabetic retinopathy or macular degeneration), an inability to reliably perform visual field testing, or likely inability to obtain images of sufficient quality from Optical Coherence Tomography (OCT) (e.g., because of severe cataract). All testing protocols were approved by the Legacy Health Institutional Review Board and adhered to the tenets of the Declaration of Helsinki.

SAP was conducted using a Humphrey Field Analyzer IIi (Carl Zeiss Meditec Inc., Dublin, CA, USA), with the SITA Standard testing strategy and 24-2 test pattern. Functional status on each visit was summarized by the average pointwise total deviation (TD), after transforming each value onto a linear 1/Lambert scale by the equation $TD_{Lin} = 10^{(TD/10)}$ to make the structure-function relation approximately linear over the range of observed values, and is denoted by $AveTD_{Lin}$.[12,13] OCT testing was performed using a Spectralis instrument (Heidelberg Engineering GmbH, Heidelberg, Germany), with a 6° radius circle scan centered on the optic nerve head to measure the average RNFLT in micrometers; automated layer segmentations were manually corrected if necessary to correct obvious errors.[14]

Participants attended study visits approximately once every six months, or as close as could be scheduled. The SEM technique (described in the next section) requires that data be from discretized time points, minimizing the proportion of missing data. Therefore visits were binned into six month time periods. Visit 1 was defined as the first visit on which both SAP and OCT results were available for a given eye. Visit 2 was defined as a visit between three to nine months after Visit 1; Visit 3 was defined as a visit between nine and 15 months after Visit 1. If more than one visit occurred within a time period, the first visit in that period was used; if no visit occurred within the time period then the results for that period were treated as missing data. The rate of change between consecutive time periods was then calculated, using the exact test dates; for example the rate of functional change in Interval 1 was defined as $\Delta AveTD_{Lin}(1) = (AveTD_{Lin}(Visit2) - AveTD_{Lin}(Visit1))/(Date(Visit2) - Date(Visit1))$. Rates were available for up to 6 such time intervals per eye (i.e. based on series of up to 7 visits).

Robust estimation using SEMs requires minimizing the proportion of missing data.[15] Since the outcomes of interest in this study take the form of the rate of change during the time interval from visit n to Visit n + 1, that outcome is missing if data were not obtained at either one of those visits. Further, to avoid the potential of bias favoring structural or functional tests, data were only included from visits at which both SAP and OCT results were available. Therefore to maximize the robustness of the results, for the primary analyses, data were not excluded on the basis of not meeting the standard reliability criteria. However, a secondary analysis was also performed, using only time points at which the OCT image quality score was ≥15 and there were fewer than 20% fixation losses and 15% false-positive results for SAP. For comparison with *detectable* change, we used linear regression to determine whether each eye ever demonstrated a significantly negative rate of change of $AveTD_{Lin}$ and/or of age-corrected[16] RNFLT based on the first n visits, for $n \geq 4$.

SEM

SEMs are a Class of Statistical Models that are used to Analyze "Structural" Models,[17] including latent growth models.[18] The most relevant advantages of using SEMs for this study compared to more traditional approaches are that they allow modeling of unobservable latent variables such as the true rates of functional and structural change, they allow and adjust for correlations between predictors, and they allow the same data point to be both predicted by the status at a previous time point and be a predictor of the status at a future time point, that is, appearing on both left and right sides of different regression equations.

In this study, we are interested in whether there is a time lag between the true rates of change in $AveTD_{Lin}$ from visit n to n + 1, denoted by $\Delta AveTD_{Lin}(n)$; and the corresponding rates of change in RNFLT, denoted by $\Delta RNFLT(n)$, regardless of whether that change is outside normal limits of variability (i.e., regardless of whether the change would be detectable in clinical care). We assume that for a given eye there are true underlying rates of functional and structural loss, represented by latent variables $F(n)$ and $S(n)$, respectively, and that these rates change linearly across the series with a fixed intercept and rate per eye. Thus the model allows the rate of progression to increase or decrease, but it does so consistently such that $^{dF}/_{dn}$ and $^{dS}/_{dn}$ remain constant; this assumption was believed to be reasonable for series of up to three years (seven visits).[19] These underlying rates $F(n)$ and $S(n)$ are assumed to be positively correlated between eyes, but they are not constrained to be proportional because there is substantial interindividual variability in the relation between structure and function even in healthy eyes.[12] The observed rate $\Delta AveTD_{Lin}(n)$ is treated as a random variable predicted by $F(n)$ but with variance σ_F^2. Similarly, the observed rate $\Delta RNFLT(n)$ is treated as a random variable predicted by $S(n)$ with variance σ_S^2. The variances σ_F^2 and σ_S^2 are assumed to be constant throughout the series, implicitly making the simplifying assumption that even though test-retest variability in perimetry is known to vary with severity, the rate of change is uncorrelated with both.

Four SEM models were constructed, differing only in their use of information from the other testing modality, and fit independently of each other:

$$\text{Model A: } \Delta AveTD_{Lin}(n) = F(n) + \alpha_A * \Delta AveTD_{Lin}(n-1) + \beta_A * \Delta RNFLT(n) + \varepsilon_F$$

$$\Delta RNFLT(n) = S(n) + \gamma_A * \Delta RNFLT(n-1) + \varepsilon_S$$

$$\text{Model B: } \Delta AveTD_{Lin}(n) = F(n) + \alpha_B * \Delta AveTD_{Lin}(n-1) + \beta_B * \Delta RNFLT(n-1) + \varepsilon_F$$

$$\Delta RNFLT(n) = S(n) + \gamma_B * \Delta RNFLT(n-1) + \varepsilon_S$$

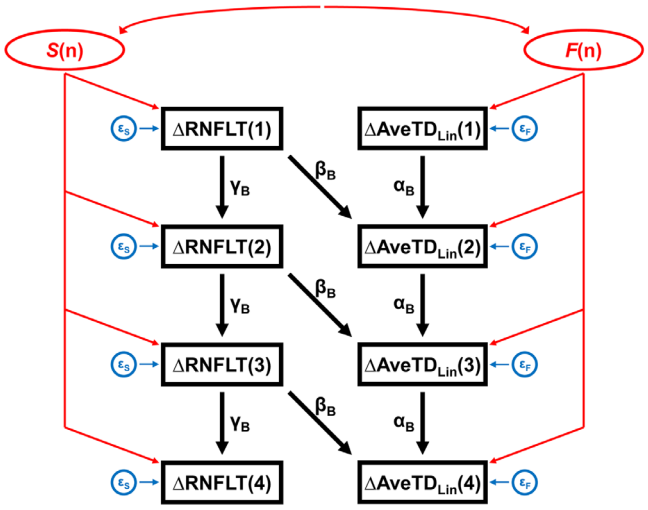

Figure 1. The path diagram for one of the four structural equation models used, Model B. The latent variables *S*(n) and *F*(n) are shown in *red*, representing the underlying rates of structural and functional change, respectively. These are connected by a *double-headed arrow*, signifying that they are assumed to be correlated. Both *S*(n) and *F*(n) change linearly with the visit number n. Series actually extended as far as $n = 6$ for some eyes. The observed variables ΔRNFLT(n) and $\Delta\text{AveTD}_{\text{Lin}}$(n) are shown in *black*, representing the measured rates of change of retinal nerve fiber layer thickness (RNFLT) and of mean linearized total deviation ($\text{AveTD}_{\text{Lin}}$) over period n (from visit n to visit n + 1). The measurement errors ε_S and ε_F are shown in *blue* and are assumed to be independent identically distributed random variables with mean zero and standard deviations σ_S and σ_F, respectively. *Directional arrows* indicate regressions, with labelled coefficients. Hence, for example, $\Delta\text{AveTD}_{\text{Lin}}(2) = F(2) + \alpha_B{*}\Delta\text{AveTD}_{\text{Lin}}(1) + \beta_B{*}\Delta\text{RNFLT}(1) + \varepsilon_F$. If coefficient β_B is positive and statistically significant, then that implies that ΔRNFLT in the previous time period is predictive of $\Delta\text{AveTD}_{\text{Lin}}$ in the current time period; that is, a time lag whereby structural change occurred earlier than and was predictive of functional change.

$$\text{Model C: } \Delta\text{RNFLT(n)} = S(n) + \alpha_C{*}\Delta\text{RNFLT(n} - 1) + \beta_C{*}\Delta\text{AveTD}_{\text{Lin}}(n) + \varepsilon_S$$

$$\Delta\text{AveTD}_{\text{Lin}}(n) = F(n) + \gamma_C{*}\Delta\text{AveTD}_{\text{Lin}}(n-1) + \varepsilon_F$$

$$\text{Model D: } \Delta\text{RNFLT(n)} = S(n) + \alpha_D{*}\Delta\text{RNFLT(n} - 1) + \beta_D{*}\Delta\text{AveTD}_{\text{Lin}}(n-1) + \varepsilon_S$$

$$\Delta\text{AveTD}_{\text{Lin}}(n) = F(n) + \gamma_D{*}\Delta\text{AveTD}_{\text{Lin}}(n-1) + \varepsilon_F$$

The path diagram for Model B, as an example, is shown in Figure 1. The error term ε_S is normally distributed with mean zero and standard deviation σ_S; and the error term ε_S is normally distributed with mean zero and standard deviation σ_F. Thus, in Model A, the rate of functional change in interval n can be predicted based on knowledge of the rate of functional change in interval n − 1, together with the rate of structural change in interval n; and there is no assumption of time lag between $\Delta\text{AveTD}_{\text{Lin}}$ and ΔRNFLT. Similarly in Model C, the rate of structural change in interval n can be predicted based on its rate in interval n − 1 and the rate of functional change in interval n, with no assumption of a time lag. In Model B, the rate of structural change in interval n − 1 helps predict the rate of functional change in interval n, i.e. there is a time lag whereby structural change occurs earlier than, and is predictive of, functional change over the next time interval. Conversely in Model D, there is a time lag whereby functional change occurs earlier than and is predictive of structural change over the next time interval.

The coefficients α, β, and γ can differ between the four models. However, coefficients α and γ are always expected to be negative; the change $\Delta\text{AveTD}_{\text{Lin}}$(n − 1) from visit n − 1 to visit n will be inversely correlated with the change $\Delta\text{AveTD}_{\text{Lin}}$(n) from visit n to visit n + 1 since they both have the measurement $\text{AveTD}_{\text{Lin}}$ at visit n in common. Coefficient β is constrained to be nonnegative, for reasons of clinical plausibility. The primary hypothesis being tested is that more rapid change in one modality may predict more rapid change in the other modality, either in the same interval or the following interval; that is, whether β is significantly greater than zero.

Analyses were performed using R statistical software, version 4.0.0,[20] with the lavaan package.[21] Models were fit using full information maximum likelihood estimation to ensure that the results are statistically consistent and unbiased despite the presence of missing data.[15] Goodness of fit for each model was assessed using the root mean square error of approximation (RMSEA)[22,23]; the Tucker Lewis index (TLI, also known as the nonnormed fit index) representing the magnitude of the improvement in fit over a null model[24]; and the adjusted goodness of fit (AGFI) representing the proportion of variance explained by the model, analogously to an adjusted R^2 value.[24]

Localized Analysis

The analyses described above used global indexes for both function ($\text{AveTD}_{\text{Lin}}$) and structure (RNFLT). These are less sensitive to the smaller localized defects that are typically seen in early glaucoma. Additionally, there is not a perfect mapping between the two. Due to the layout of test locations in the 24-2 grid, $\text{AveTD}_{\text{Lin}}$ will better detect changes in the superior and inferior mid-peripheral regions where there are several test locations, versus in the central or temporal regions where there are very few test locations. By contrast, RNFLT is a simple mean value around the circumpapillary scan, and so gives equal weight to defects occurring at all angles around the optic nerve head. Therefore two localized analyses were also performed. RNFLT(Sup) was defined as the average RNFLT within the combined superior nasal and superior temporal 40°-wide sectors that are output by the Spectralis OCT software. This was compared against $\text{AveTD}_{\text{Lin}}$(Inf), defined as the average of the sensitivities (on the linear scale as before) of the corresponding 21 visual field locations, based on the map of Garway-Heath et al.[25]; namely all locations in the inferior hemifield, except for the three locations closest to the blind spot at 3° below the horizontal midline and the two locations temporal of the blind spot. Similarly, RNFLT(Inf) was defined as the average RNFLT within the combined inferior nasal and inferior temporal 40°-wide sectors, and was compared against $\text{AveTD}_{\text{Lin}}$(Sup) defined as the average of the corresponding 21 locations in the superior hemifield of the visual field.

Results

Table 1 summarizes the cohort. An available interval was defined as two visits occurring in consecutive six-month time periods; for example, one visit occurring nine to 15 months after their first visit, and another visit occurring

Table 1. Characteristics of the Dataset Used

	Mean	Standard Deviation	Range
Age at first visit (years)	68.8	9.3	44.9 to 90.3
Initial mean deviation (dB)	−1.22	3.3	−20.4 to +3.0
Initial retinal nerve fiber layer thickness (µm)	80.2	14.8	36.5 to 114.2
Number of available intervals	3.6	1.7	1 to 6
Number of available reliable intervals	2.8	1.8	0 to 6

An "Available Interval" is defined as the availability of data from two visits in consecutive six-month time periods; an "Available Reliable Interval" also requires that results from both visits met the manufacturer's recommended reliability criteria.

15 to 21 months after their first visit. Data were available for a total of 1135 intervals, from 318 eyes of 164 participants. The mean time between the two visits for an interval was 207 days, ranging from 161 to 364 days. Thirty of the eyes demonstrated a significantly negative rate of change of mean deviation based on visits 1-n, for some number of visits $n \geq 4$, by the end of their series; 57 eyes demonstrated a significantly negative rate of age-corrected RNFLT before the end of their series (comparison $P = 0.001$, McNemar's test).

The fit coefficients for the first three SEM models were as follow:

$$\text{Model A: } \Delta\text{AveTD}_{\text{Lin}}(n) = F(n) - 0.550 * \Delta\text{AveTD}_{\text{Lin}}(n-1) + 0 * \Delta\text{RNFLT}(n) + \varepsilon_F$$

$$\Delta\text{RNFLT}(n) = S(n) - 0.313 * \Delta\text{RNFLT}(n-1) + \varepsilon_S$$

$$\text{Model B: } \Delta\text{AveTD}_{\text{Lin}}(n) = F(n) - 0.550 * \Delta\text{AveTD}_{\text{Lin}}(n-1) + 0 * \Delta\text{RNFLT}(n-1) + \varepsilon_F$$

$$\Delta\text{RNFLT}(n) = S(n) - 0.313 * \Delta\text{RNFLT}(n-1) + \varepsilon_S$$

$$\text{Model C: } \Delta\text{RNFLT}(n) = S(n) - 0.313 * \Delta\text{RNFLT}(n-1) + 0 * \Delta\text{nAveTD}_{\text{Lin}}(n) + \varepsilon_S$$

$$\Delta\text{AveTD}_{\text{Lin}}(n) = F(n) - 0.550 * \Delta\text{AveTD}_{\text{Lin}}(n-1) + \varepsilon_F$$

That is, after adjusting for the rate of functional change during the previous interval, neither the rates of structural change ΔRNFLT in the same interval (Model A) nor the previous interval (Model B) were significant predictors of the rate of functional change. Similarly, after adjusting for the rate of structural change during the previous interval, the rate of functional change $\Delta\text{AveTD}_{\text{Lin}}$ in the same interval (Model C) was not a significant predictor of the rate of structural change. Thus the fitted coefficients for Model A, Model B, and Model C are all identical, as seen in Figure 2.

The latent variable $F(n)$ had an intercept of $-0.043\ \text{L}^{-1}\text{y}^{-1}$ in Models A-C, indicating the predicted rate of change in $\text{AveTD}_{\text{Lin}}$ within an interval if the rate in the previous interval had been zero. Similarly, the latent variable $S(n)$ had intercept −0.717 µm/y, indicating the rate of change in RNFLT if the rate in the previous interval had been zero. It should be noted at this point that $\text{AveTD}_{\text{Lin}}$ is age-corrected whereas RNFLT is not; the only effect of this inconsistency is to alter these constant intercepts for $S(n)$ and $F(n)$, and age correction would not alter the magnitude or statistical significance of any of the other coefficients. As predicted, the fitted values of coefficients α and γ were both negative, because of the effect of variability in the measurement at time n on the rates $\Delta\text{AveTD}_{\text{Lin}}(n-1)$ and $\Delta\text{AveTD}_{\text{Lin}}(n)$ and similarly for

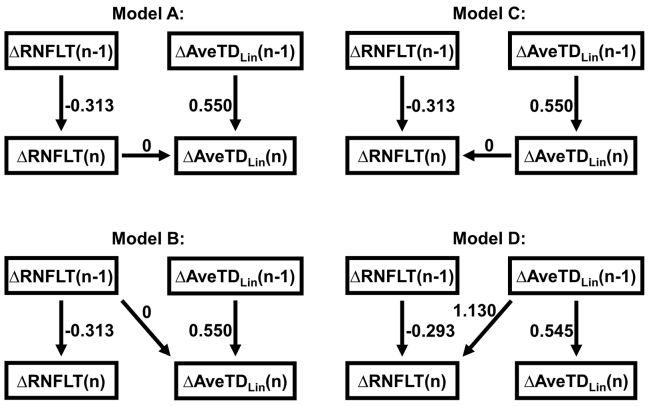

Figure 2. Fitted coefficients for the four structural equation models used. Only coefficients relating the observed variables are shown, representing the measured rates of change of RNFLT and of $\text{AveTD}_{\text{Lin}}$ over period n (from visit n to visit n + 1).

RNFLT. If there were no true changes in $\text{AveTD}_{\text{Lin}}$ or RNFLT, we would expect both α and γ to be −1; that is, the average rate of change in interval n would be the exact opposite of the rate in interval n − 1. The fact that both α and γ are greater than −1 indicates that the true rates of change in intervals n − 1 and n would be correlated in the absence of measurement variability. It should also be noted that the coefficient β was constrained to be nonnegative to maintain physiological plausibility; hence, the best fit estimate is $\beta = 0$ for all three of the above models. Removing that constraint gave slightly negative but nonsignificant values for β_A ($P = 0.520$) and β_B ($P = 0.062$). The unconstrained fitted value of β_C was −0.909 with $P = 0.017$.

However, for the fourth SEM model:

$$\text{Model D: } \Delta\text{RNFLT}(n) = S(n) - 0.293 * \Delta\text{RNFLT}(n-1) + 1.130 * \Delta\text{AveTD}_{\text{Lin}}(n-1) + \varepsilon_S$$

$$\Delta\text{AveTD}_{\text{Lin}}(n) = F(n) - 0.545 * \Delta\text{AveTD}_{\text{Lin}}(n-1) + \varepsilon_F$$

The rate of functional change in the previous interval was a significant predictor of the rate of structural change in the current interval, with $P = 0.007$, and 95% confidence interval (CI) for the coefficient (0.308, 1.953). That is, more rapid structural change in interval n − 1 did not predict more rapid functional change in interval n (Model B); but more rapid functional change in interval n − 1 did predict more rapid structural change in interval n (Model D). Overall, this suggests that functional change as measured by $\text{AveTD}_{\text{Lin}}$ predicts and precedes structural change as measured by RNFLT.

Table 2. Goodness-of-Fit Measures for the Models

Model	RMSEA	TLI	AGFI
All visits			
Model A	0.107 (0.094–0.120)	0.563	68.0%
Model B	0.107 (0.094–0.120)	0.563	68.0%
Model C	0.107 (0.094–0.120)	0.563	68.0%
Model D	0.105 (0.093–0.118)	0.577	68.3%
Reliable visits only			
Model A	0.114 (0.101–0.129)	0.487	64.5%
Model B	0.114 (0.101–0.129)	0.487	64.5%
Model C	0.114 (0.101–0.129)	0.487	64.5%
Model D	0.112 (0.098–0.126)	0.508	65.3%

In all, 88% of individual visits were considered reliable; that is, there were ≤20% fixation losses and ≤15% false-positive results for SAP, and an OCT scan quality of ≥15. The regression equations in the SEM models require that data be available for two consecutive intervals, that is, for test dates falling within three consecutive six-month time periods. There were 743 pairs of consecutive intervals but just 538 pairs of consecutive reliable intervals. Thus, even though 88% of individual visits were considered reliable, only 72% of the available consecutive intervals could be used for the secondary analysis restricted to reliable data only. The best fit coefficients using only reliable data were as follow:

$$\text{Model A: } \Delta\text{AveTD}_{\text{Lin}}(n) = F(n) - 0.616 * \Delta\text{AveTD}_{\text{Lin}}(n-1) + 0 * \Delta\text{RNFLT}(n) + \varepsilon_F$$

$$\Delta\text{RNFLT}(n) = S(n) - 0.346 * \Delta\text{RNFLT}(n-1) + \varepsilon_S$$

$$\text{Model B: } \Delta\text{AveTD}_{\text{Lin}}(n) = F(n) - 0.616 * \Delta\text{AveTD}_{\text{Lin}}(n-1) + 0.001 * \Delta\text{RNFLT}(n-1) + \varepsilon_F$$

$$\Delta\text{RNFLT}(n) = S(n) - 0.346 * \Delta\text{RNFLT}(n-1) + \varepsilon_S$$

$$\text{Model C: } \Delta\text{RNFLT}(n) = S(n) - 0.346 * \Delta\text{RNFLT}(n-1) + 0 * \Delta\text{AveTD}_{\text{Lin}}(n) + \varepsilon_S$$

$$\Delta\text{AveTD}_{\text{Lin}}(n) = F(n) - 0.616 * \Delta\text{AveTD}_{\text{Lin}}(n-1) + \varepsilon_F$$

$$\text{Model D: } \Delta\text{RNFLT}(n) = S(n) - 0.315 * \Delta\text{RNFLT}(n-1) + 1.634 * \Delta\text{AveTD}_{\text{Lin}}(n-1) + \varepsilon_S$$

$$\Delta\text{AveTD}_{\text{Lin}}(n) = F(n) - 0.608 * \Delta\text{AveTD}_{\text{Lin}}(n-1) + \varepsilon_F$$

All four models are qualitatively similar to those derived including unreliable test visits. In Model D, the coefficient β_D for $\Delta\text{AveTD}_{\text{Lin}}(n-1)$ now has $P = 0.003$ and 95% CI for the coefficient (0.541, 2.727), supporting the robustness of the model.

Table 2 shows a selection of goodness-of fit measures for each model.[24] Model D fit the data slightly better than Models A to C, as indicated by lower RMSEA (lower prediction error), higher TLI (greater improvement in fit over a null model), and higher AGFI (greater proportion of variance explained). Although direct comparisons should be treated with caution, the models using only reliable data appeared to have slightly worse fits to the data, presumably because of reduced sample size.

When using localized analyses, results were similar. In Models A, B, and C, coefficients β_A, β_B, and β_C were zero when using RNFLT(Sup) against $\Delta\text{AveTD}_{\text{Lin}}$(Inf) and when using RNFLT(Inf) against $\Delta\text{AveTD}_{\text{Lin}}$(Sup), regardless of whether only reliable visits were used. However in Model D, coefficient β_D was consistently greater than zero. Using data from all visits:

Inferior hemifield:

$$\Delta\text{RNFLT(Sup)}(n) = S(n) - 0.366 * \Delta\text{RNFLT(Sup)}(n-1) + 1.651 * \Delta\text{AveTD}_{\text{Lin}}\text{(Inf)}(n-1) + \varepsilon_S$$

$$\Delta\text{AveTD}_{\text{Lin}}\text{(Inf)}(n) = F(n) - 0.541 * \Delta\text{AveTD}_{\text{Lin}}\text{(Inf)}(n-1) + \varepsilon_F$$

Superior hemifield:

$$\Delta\text{RNFLT(Inf)}(n) = S(n) - 0.392 * \Delta\text{RNFLT(Inf)}(n-1) + 0.450 * \Delta\text{AveTD}_{\text{Lin}}\text{(Sup)}(n-1) + \varepsilon_S$$

$$\Delta\text{AveTD}_{\text{Lin}}\text{(Sup)}(n) = F(n) - 0.529 * \Delta\text{AveTD}_{\text{Lin}}\text{(Sup)}(n-1) + \varepsilon_F$$

The coefficients β_D had $P = 0.022$ for $\Delta\text{AveTD}_{\text{Lin}}$(Inf), and $P = 0.356$ for $\Delta\text{AveTD}_{\text{Lin}}$(Sup). Although those coefficients appear quite different from one another, the 95% CIs were wide; 0.236 to 3.066 for $\Delta\text{AveTD}_{\text{Lin}}$(Inf), and −0.51 to 1.405 for $\Delta\text{AveTD}_{\text{Lin}}$(Sup).

Using data from reliable visits only:

Inferior hemifield:

$$\Delta\text{RNFLT(Sup)}(n) = S(n) - 0.331 * \Delta\text{RNFLT(Sup)}(n-1) + 2.545 * \Delta\text{AveTD}_{\text{Lin}}\text{(Inf)}(n-1) + \varepsilon_S$$

$$\Delta\text{AveTD}_{\text{Lin}}\text{(Inf)}(n) = F(n) - 0.653 * \Delta\text{AveTD}_{\text{Lin}}\text{(Inf)}(n-1) + \varepsilon_F$$

Superior hemifield:

$$\Delta\text{RNFLT(Inf)}(n) = S(n) - 0.428 * \Delta\text{RNFLT(Inf)}(n-1) + 1.316 * \Delta\text{AveTD}_{\text{Lin}}\text{(Sup)}(n-1) + \varepsilon_S$$

$$\Delta\text{AveTD}_{\text{Lin}}\text{(Sup)}(n) = F(n) - 0.567 * \Delta\text{AveTD}_{\text{Lin}}\text{(Sup)}(n-1) + \varepsilon_F$$

In this case, the coefficients β_D had p = 0.006 for $\Delta\text{AveTD}_{\text{Lin}}$(Inf), and $P = 0.032$ for $\Delta\text{AveTD}_{\text{Lin}}$(Sup).

Discussion

In this study, we used SEMs to investigate whether there is a time lag between the *true* changes in structural and functional measures (RNFLT and SAP, respectively). Because of their ability to incorporate latent variables in time series for which the same observed value is both dependent on previous values and predictive of future values, SEMs allow us to test the relative utility of time-lagged variables and, hence, to identify time lags between the true rates of change even when those changes do not exceed normal variability (i.e., they are not "detectable" clinically). We found that the rate of functional change in a given time interval was predictive of the rate of structural change in the following time interval, but the converse was not true. This is despite the fact that more eyes demonstrated *detectable* change for RNFLT than for SAP, when defined as a significantly negative rate of change over ≥4 visits. This time lag implies that although *detectable change* may occur sooner for RNFLT than for SAP, *true change* for SAP occurs sooner than and is predictive of subsequent change in RNFLT.

Several previous studies have suggested that glaucomatous damage may be able to be *detected* sooner using structural testing than functional testing. A classic study by Kerrigan-Baumrind et al.[1] is often cited as evidence that 25% to 35% of retinal ganglion cells (RGCs) are lost before functional abnormalities develop that are measurable using SAP; although that interpretation of their results has been questioned[26] because their results actually showed a 6-decibel (dB) functional loss in eyes without RGC loss.[27] A more recent study by Wollstein et al.[2] suggested that ~17% of RNFLT, as measured by optical coherence tomography (OCT), must be lost before functional abnormalities were detected; although it should be noted that due to the specific purpose of their study that conclusion was based on a segmented linear regression fit on decibel-scaled sensitivity values rather than the more commonly-used linear fit on linear-transformed sensitivities. A study by Kuang et al.[28] also found evidence of RNFL defects before confirmed functional loss, suggesting higher sensitivity, although they did not look for functional loss being detected before RNFL thinning. Progression was more commonly detected using optic disc photos before perimetry in the OHTS (3.0% of eyes, vs. 1.7% of eyes for which progression was detected the other way round)[4]; but this in part reflects conformational changes to the optic nerve head connective tissues that do not affect peripapillary RNFL measurements.[29] These comparisons are also heavily influenced by the greater test-retest variability of SAP,[5,7] which means that even if both structure and function change concurrently by the same amount, that change would exceed normal limits sooner for structural change.[5]

It is not surprising that true functional change could occur earlier than RNFL thinning. The largest proportion of the RNFL consists of axons. These axons likely stop being fully functional some time before their soma undergo apoptosis and efferocytosis. This unknown time gap may be short, but it must be present, because vision function cannot persist after an RGC or its axon are gone. The only way RNFL thinning could precede, and be predictive of, subsequent functional loss would be if individual axons or non-neural tissue became thinner because of compression, stretching, or both. This transverse compression has been reported in rim tissue in experimental glaucoma, but it was not detected at the 6° distance from the centroid of the optic nerve head at which the circle scan used to measure RNFLT is conducted.[30] On the contrary, evidence suggests that RNFL thinning is actually significantly delayed after axon loss,[8,9] postulated to be due to a combination of axon enlargement[31] and mechanical resistance to layer collapse,[32] causing a decreased density of axons within the layer but little or no change in the layer thickness, something that has also been suggested in human glaucoma based on results from adaptive optics scanning laser ophthalmoscopy.[33] Furthermore, there is evidence of RGC functional deficits preceding loss of RGCs in experimental models of glaucoma, such as impaired axonal transport[34] and altered neural responsiveness to stimuli.[35–38] Thus results from animal models suggest functional loss preceding RGC loss preceding RNFL thinning.

Our results extend that principle to human glaucoma, showing not only that RNFL thinning lags behind functional loss, but also that the time lag is substantial enough to be meaningful. Longitudinal testing in this study was performed once every six months. This precludes providing a highly precise estimate of the time lag, which might vary substantially between eyes anyway. In further analyses (not shown), there was no evidence supporting a time lag of one year (i.e., two time intervals), but the sample size was much smaller for those analyses, and so we cannot definitively conclude that the duration of the time lag is only one time interval. However, we can reasonably conclude that the average time lag is on average closer to six months than it is to zero.

Use of the SEM framework in this study reveals a time lag that may not be detected using simpler statistical techniques, but also comes with inherent disadvantages. One is that the models require use of data collected at discrete time intervals. Thus a major advantage of our study population in this context is the regularity of the testing interval; 72% of intervisit intervals were within ±30 days of six months, and 85% were within ±60 days. Additionally, the shortest actual time interval was 161 days, and so there are none of the unrealistically large rates of change of structure or function that could occur if tests happened only a few days apart.

A bigger caveat with the results is the unusual decision to include study visits with unreliable test results to minimize the proportion of missing data. Only 0.9% of OCT images had quality score <15, the manufacturer's recommended cutoff. Of the visual fields, 0.9% had >15% false-positive results. A far higher proportion of visual fields, 16.0%, had >20% fixation losses; most of those tests did in fact have subjectively acceptable fixation as viewed by the technician on the instrument's monitor, with the recorded level of fixation losses reflecting inaccurate mapping of the blind spot in the initial phase of the test. Moreover, the effect of these unreliable results can safely be assumed to just be an increase in test-retest variability, rather than imparting any bias on the results, because they are distributed randomly through a patient's test series. Indeed, the goodness-of-fit measures reported in Table 2 were consistently slightly worse when restricting the data to reliable tests only, possibly because of the reduced sample size. Excluding data from a single visit could reduce the number of available pairs of consecutive intervals by three. Given the reduction in power of SEMs caused by missing data,[15] even in a case such as this where the data can be considered as missing completely at random, it was felt preferable to include all data for the primary analysis. Reassuringly, the secondary analysis restricted to only reliable test results gave very similar model fits. The coefficients were of slightly larger magnitude, indicating that the data from the previous time period or other modality were slightly more predictive due to being more reliable, but differences were small.

The primary analysis used global measures of both function and structure, which have lower variability due to averaging out much of the measurement noise. Results were very similar when using localized analyses, which compared the average sensitivities at 21 visual field locations against the RNFLT within the corresponding 80° sector, strengthening our confidence in the conclusions. With our current data, the higher variability precludes performing more localized or pointwise analyses, but these would be of interest once more data is available.

Even with global measures, the goodness-of-fit indexes reported in Table 2 are weaker than those typically recommended when using SEMs. Commonly-used definitions of a "good fit" would be a TLI > 0.95, or AGFI > 0.90; various cutoffs for RMSEA have been suggested from <0.05 to <0.10.[24] However, use of these kinds of fixed cutoffs has been criticized, because it does not take into account the nature of the data or of the research question.[22] In particular, AGFI reflects the percentage of variance explained, adjusted for the number of free parameters, analogous to

the more familiar adjusted R^2 for regression models. Given the well-known variability of perimetric measures in particular,[39,40] a method that can explain 60% to 70% of the variance in the rate of change between two visits should probably be considered as an impressively good fit to the data; the standard SEM cutoff of explaining 90% of the variance is unrealistic given the level of measurement variability.

The conclusions drawn are specific to the test modalities used. In particular, minimum rim width (MRW) may show changes sooner than RNFLT.[29] It could be hypothesized that the observed time lag may not be present or may even be reversed, when considering MRW versus SAP instead of RNFLT versus SAP.[30]

SAP sensitivities were transformed onto a linear 1/Lambert scale before averaging, because this has been reported as resulting in a more linear structure-function relation.[12,13] However, the true relation may be better represented by a nonlinear or segmented linear fit. The Size III stimulus is larger than Ricco's area in early glaucoma, potentially causing SAP to underestimate RGC loss and hence RNFL thickness; whereas at more damaged locations Ricco's area expands and exceeds this stimulus size.[41–43] For this study, information was averaged across multiple locations, often including some locations with sensitivities above this break point and others below, and so the impact of this caveat is hard to predict; but it seems unlikely that it would cause functional changes to artefactually appear to precede structural changes as seen in our results.

A final caveat is that the cohort consists mostly of patients with early glaucoma or glaucoma suspects. Furthermore, both eyes are tested even if only one eye has developed glaucomatous damage, on the basis that the fellow eye is at increased risk of developing glaucoma.[44,45] Indeed, only 19% of eyes had MD worse than −3 dB at the start of their series, and 8% worse than −6 dB. The presence and magnitude of the time lag between changes in function and in RNFLT may vary through the course of the disease process.

In summary, we present evidence of a time lag of several months, whereby loss of visual function precedes and is predictive of thinning of the RNFL in human glaucoma. This does not negate previous findings suggesting that change in RNFLT may be *detectable* sooner using current testing techniques, because of lower variability, meaning that less change is needed to be outside normal limits. However, the presence of a time lag should be taken into account when considering the pathophysiologic mechanisms that cause glaucomatous damage. It also encourages the development of improved and less variable functional testing, and use of alternative structural measures, which may allow earlier detection of damage and provide better prognostic information about disease progression.

Acknowledgments

Supported by NIH NEI Grant R01-EY020922 (SKG); unrestricted research support from The Legacy Good Samaritan Foundation, Portland, OR, USA.

Disclosure: **S.K. Gardiner,** Heidelberg Engineering (F, R); **S.L. Mansberger,** Abbvie (F), Thea Pharmaceuticals (C), Nicox Pharmaceuticals (C); **B. Fortune,** Heidelberg Engineering (F), Perfuse Therapeutics Inc. (F, R)

References

1. Kerrigan-Baumrind LA, Quigley HA, Pease ME, Kerrigan DF, Mitchell RS. Number of ganglion cells in glaucoma eyes compared with threshold visual field tests in the same persons. *Invest Ophthalmol Vis Sci*. 2000;41:741–748.
2. Wollstein G, Kagemann L, Bilonick RA, et al. Retinal nerve fibre layer and visual function loss in glaucoma: the tipping point. *Br J Ophthalmol*. 2012;96:47–52.
3. Wollstein G, Schuman JS, Price LL, et al. Optical coherence tomography longitudinal evaluation of retinal nerve fiber layer thickness in glaucoma. *Arch Ophthalmol*. 2005;123:464–470.
4. Keltner JL, Johnson CA, Anderson DR, et al. The association between glaucomatous visual fields and optic nerve head features in the Ocular Hypertension Treatment Study. *Ophthalmology*. 2006;113:1603–1612.
5. Gardiner SK, Fortune B, Demirel S. Signal-to-Noise Ratios for Structural and Functional Tests in Glaucoma. *Transl Vis Sci Technol*. 2013;2:3.
6. Weinreb RN, Garway-Heath DF, Leung C, Medeiros FA, Liebmann JM. Structure Function. *Diagnosis of Primary Open Angle Glaucoma: WGA consensus series* - 10: Kugler Publications. 2017:90–127.
7. Gardiner SK, Johnson CA, Demirel S. The effect of test variability on the structure-function relationship in early glaucoma. *Graefes Arch Clin Exp Ophthalmol*. 2012;250:1851–1861.
8. Cull GA, Reynaud J, Wang L, Cioffi GA, Burgoyne CF, Fortune B. Relationship between orbital optic nerve axon counts and retinal nerve fiber layer thickness measured by spectral domain optical coherence tomography. *Invest Ophthalmol Vis Sci*. 2012;53:7766–7773.
9. Fortune B, Hardin C, Reynaud J, et al. Comparing Optic Nerve Head Rim Width, Rim Area, and Peripapillary Retinal Nerve Fiber Layer Thickness to Axon Count in Experimental Glaucoma. *Invest Ophthalmol Vis Sci*. 2016;57:OCT404–OCT412.
10. Gardiner SK, Mansberger SL, Demirel S. Detection of Functional Change Using Cluster Trend Analysis in Glaucoma. *Invest Ophthalmol Vis Sci*. 2017;58:Bio180–Bio190.
11. Gardiner SK, Fortune B, Demirel S. Localized Changes in Retinal Nerve Fiber Layer Thickness as a Predictor of Localized Functional Change in Glaucoma. *Am J Ophthalmol*. 2016;170:75–82.
12. Hood DC, Anderson SC, Wall M, Kardon RH. Structure versus function in glaucoma: an application of a linear model. *Invest Ophthalmol Vis Sci*. 2007;48:3662–3668.
13. Harwerth RS, Wheat JL, Fredette MJ, Anderson DR. Linking structure and function in glaucoma. *Prog Retin Eye Res*. 2010;29:249–271.
14. Mansberger SL, Menda SA, Fortune BA, Gardiner SK, Demirel S. Automated Segmentation Errors When Using Optical Coherence Tomography to Measure Retinal Nerve Fiber Layer Thickness in Glaucoma. *American Journal of Ophthalmology*. 2017;174:1–8.
15. Allison PD. Missing data techniques for structural equation modeling. *Journal of abnormal psychology*. 2003;112:545.
16. Chauhan BC, Danthurebandara VM, Sharpe GP, et al. Bruch's Membrane Opening Minimum Rim Width and Retinal Nerve Fiber Layer Thickness in a Normal White Population: A Multicenter Study. *Ophthalmology*. 2015;122:1786–1794.
17. Kaplan D. *Structural equation modeling: Foundations and extensions*. Thousand Oaks, CA: Sage; 2000.
18. Meredith W, Tisak J. Latent curve analysis. *Psychometrika*. 1990;55:107–122.
19. 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;54:1252–1259.

20. R Development Core Team. *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. 2013.
21. Rosseel Y. Lavaan: An R package for structural equation modeling and more. Version 0.5–12 (BETA). *Journal of Statistical Software*. 2012;48:1–36.
22. Chen F, Curran PJ, Bollen KA, Kirby J, Paxton P. An Empirical Evaluation of the Use of Fixed Cutoff Points in RMSEA Test Statistic in Structural Equation Models. *Sociol Methods Res*. 2008;36:462–494.
23. Cangur S, Ercan I. Comparison of model fit indices used in structural equation modeling under multivariate normality. *Journal of Modern Applied Statistical Methods*. 2015;14:14.
24. Schermelleh-Engel K, Moosbrugger H, Müller H. Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research Online*. 2003;8:23–74.
25. Garway-Heath DF, Poinoosawmy D, Fitzke FW, Hitchings RA. Mapping the visual field to the optic disc in normal tension glaucoma eyes. *Ophthalmology*. 2000;107:1809–1815.
26. Malik R, Swanson WH, Garway-Heath DF. 'Structure-function relationship' in glaucoma: past thinking and current concepts. *Clinical & Experimental Ophthalmology*. 2012;40:369–380.
27. Hood DC. Does Retinal Ganglion Cell Loss Precede Visual Field Loss in Glaucoma? *Journal of Glaucoma*. 2019;28:945–951.
28. Kuang TM, Zhang C, Zangwill LM, Weinreb RN, Medeiros FA. Estimating Lead Time Gained by Optical Coherence Tomography in Detecting Glaucoma before Development of Visual Field Defects. *Ophthalmology*. 2015;122:2002–2009.
29. Chauhan BC, O'Leary N, AlMobarak FA, et al. Enhanced Detection of Open-angle Glaucoma with an Anatomically Accurate Optical Coherence Tomography-Derived Neuroretinal Rim Parameter. *Ophthalmology*. 2013;120:535–543.
30. Fortune B, Reynaud J, Hardin C, Wang L, Sigal IA, Burgoyne CF. Experimental Glaucoma Causes Optic Nerve Head Neural Rim Tissue Compression: A Potentially Important Mechanism of Axon Injury. *Investigative Ophthalmology & Visual Science*. 2016;57:4403–4411.
31. Zhu Y, Pappas AC, Wang R, Seifert P, Sun D, Jakobs TC. Ultrastructural Morphology of the Optic Nerve Head in Aged and Glaucomatous Mice. *Invest Ophthalmol Vis Sci*. 2018;59:3984–3996.
32. Fortune B. Pulling and Tugging on the Retina: Mechanical Impact of Glaucoma Beyond the Optic Nerve Head. *Investigative Ophthalmology & Visual Science*. 2019;60:26–35.
33. Hood DC, Chen MF, Lee D, et al. Confocal Adaptive Optics Imaging of Peripapillary Nerve Fiber Bundles: Implications for Glaucomatous Damage Seen on Circumpapillary OCT Scans. *Transl Vis Sci Technol*. 2015;4:12.
34. Fahy ET, Chrysostomou V, Crowston JG. Mini-Review: Impaired Axonal Transport and Glaucoma. *Current Eye Research*. 2016;41:273–283.
35. Weber AJ, Harman CD. Structure-function relations of parasol cells in the normal and glaucomatous primate retina. *Invest Ophthalmol Vis Sci*. 2005;46:3197–3207.
36. Harwerth RS, Crawford ML, Frishman LJ, Viswanathan S, Smith EL, 3rd, Carter-Dawson L. Visual field defects and neural losses from experimental glaucoma. *Prog Retin Eye Res*. 2002;21:91–125.
37. Risner ML, Pasini S, Cooper ML, Lambert WS, Calkins DJ. Axogenic mechanism enhances retinal ganglion cell excitability during early progression in glaucoma. *Proceedings of the National Academy of Sciences of the United States of America*. 2018;115:E2393–E2402.
38. Risner ML, McGrady NR, Pasini S, Lambert WS, Calkins DJ. Elevated ocular pressure reduces voltage-gated sodium channel NaV1.2 protein expression in retinal ganglion cell axons. *Exp Eye Res*. 2020;190:107873.
39. Heijl A, Lindgren A, Lindgren G. Test-retest variability in glaucomatous visual fields. *Am J Ophthalmol*. 1989;108:130–135.
40. 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.
41. Redmond T, Garway-Heath DF, Zlatkova MB, Anderson RS. Sensitivity loss in early glaucoma can be mapped to an enlargement of the area of complete spatial summation. *Invest Ophthalmol Vis Sci*. 2010;51:6540–6548.
42. Khuu SK, Kalloniatis M. Standard Automated Perimetry: Determining Spatial Summation and Its Effect on Contrast Sensitivity Across the Visual Field. *Investigative Ophthalmology & Visual Science*. 2015;56:3565–3576.
43. Gardiner SK. Differences in the Relation Between Perimetric Sensitivity and Variability Between Locations Across the Visual Field. *Investigative Ophthalmology & Visual Science*. 2018;59:3667–3674.
44. Susanna R, Drance SM, Douglas GR. The visual prognosis of the fellow eye in uniocular chronic open-angle glaucoma. *Britisb Journal of Ophthalmology*. 1978;62:327–329.
45. Chen PP, Park RJ. Visual field progression in patients with initially unilateral visual field loss from chronic open-angle glaucoma. *Ophthalmology*. 2000;107:1688–1692.

Investigative Ophthalmology & Visual Science