Choice of Statistical Method Influences Apparent Association Between Structure and Function in Glaucoma

Iván Marín-Franch,^{1,2} Rizwan Malik,³ David P. Crabb,² and William H. Swanson¹

¹Indiana University School of Optometry, Bloomington, Indiana

²Department of Optometry and Visual Science, City University London, London, United Kingdom

³NIHR Biomedical Research Centre for Ophthalmology at Moorfields Eye Hospital, NHS Foundation Trust and UCL Institute of Ophthalmology, London, United Kingdom

Correspondence: Iván Marín-Franch, Indiana University School of Optometry, 800 E. Atwater Avenue, Bloomington, IN 47405; imarinfr@indiana.edu.

Submitted: June 11, 2012

Accepted: April 25, 2013

Citation: Marín-Franch I, Malik R, Crabb DP, Swanson WH. Choice of statistical method influences apparent association between structure and function in glaucoma. *Invest Ophthalmol Vis Sci.* 2013;54:4189–4196. DOI:10.1167/iovs.12-10377

PURPOSE. The aim of this study was to explore how different statistical methods may lead to inconsistent inferences about the association between structure and function in glaucoma.

METHODS. Two datasets from published studies were selected for their illustrative value. The first consisted of measurements of neuroretinal rim area in the superior-temporal sector paired with the corresponding visual field sensitivity. The second consisted of measurements of average retinal nerve fiber layer thickness over all sectors paired with the corresponding visual field sensitivity. Statistical methods included linear and segmented regression, and a nonparametric local-linear fit known as loess. The analyses were repeated with all measurements expressed as percent of mean normal.

RESULTS. Slopes from linear fits to the data changed by a factor of 10 depending on the linear regression method applied. Inferences about whether structural abnormality precedes functional abnormality varied with the statistical design and the units of measure used.

CONCLUSIONS. The apparent association between structure and function in glaucoma, and consequent interpretation, varies with the statistical method and units of measure. Awareness of the limitations of any statistical analysis is necessary to avoid finding spurious results that ultimately may lead to inadequate clinical recommendations.

Keywords: ordinary least-squares regression, segmented regression, loess, structure-function relationships

In patients with glaucoma, structural damage can be assessed with imaging devices from which the area of the neuroretinal rim area and the thickness of the retinal nerve fiber layer (RNFL) can be estimated. In addition, visual function loss can be assessed with static automated perimetry, which provides global indices concerning the visual field as well as estimates of visual sensitivities at individual locations. Since structural and functional measures are surrogates for the severity of glaucomatous damage, they should be related to each other.

Many studies of the association between structural and functional measures have been done. The purpose of these studies was to understand better the nature of the disease and to evaluate clinical measures against each other. If the association is very strong (e.g., with the coefficient of determination $R^2 > 0.9$), then one could be more confident that both measures are good surrogates for the stage of glaucoma. If the association is weak or moderate (e.g., with $R^2 < 0.5$), then one should be reluctant to accept any statistical model as a good descriptor of association, as large errors are likely to happen when predicting one measure from the other for individual subjects. Moreover, understanding the nature of the association would help clinicians to grade better the severity and stage of the disease. In experimental observations, however, the nature and strength of the association are obscured by sources of variability, including test-retest and intersubject variability, and learning and fatigue effects in perimetry, and artifacts in image acquisition such as misalign-

ments. In addition, anatomical differences between subjects result in errors in pairing structural and functional values using maps of average projections from retinal locations to sectors of the optic disc.^{1–3}

A not so often recognized difficulty in studies of association between structural and functional measures is that results depend on the choice of the statistical methods. Inferences drawn from different methods may be inconsistent with each other, and potentially misleading when the underlying assumptions are not met. Linear ordinary least-squares (OLS) regression, or simple linear regression, is the most common fitting procedure used to determine the association between structure and function. In most studies, structure is taken as the explanatory (independent) variable and function as the response (dependent) variable, and linear OLS regression of function on structure performed.^{2,4–6} An exception to this general approach was the work by Hood et al., who developed a model based on static automated perimetry and cortical evoked potentials, and tested the ability of the model to describe associations between structure and function in independent datasets.^{7,8}

Two key assumptions underpinning linear OLS regression are that the explanatory variable is observed with no measurement error and that there are no intersubject differences, in the sense that there is a one-to-one linear correspondence between the explanatory and response variables. In this type of regression, statistical fluctuations and departures from linearity, thus, are attributed to the response

variable. Attributing all variability only to function leads to biased results,⁹ unless such variability is small relative to the ranges of the “true” structural and functional values, so that R^2 is close to 1. Bias also would be present if the functional measure were taken as the explanatory variable and all variability attributed to structure, as if linear OLS were performed with the axes swapped. Many other linear regression methods exist,¹⁰ some of which acknowledge variability in both axes, but assumptions about the nature of data variability are inescapable and also are likely to lead to bias.

The aim of this study was to explore how results and conclusions about associations between structure and function in patients with glaucoma and healthy subjects vary when different statistical methods are applied to the same datasets. Methods included linear OLS regression, standardized major-axis (SMA) regression,¹¹ segmented regression,¹² and a type of nonparametric local-linear fit, known as loess,¹³ which was used to explore linearity between structural and functional methods. The intention of this study was to demonstrate how different fitting methods might lead to inconsistent inferences, rather than seek the most appropriate method of analysis, a much more arduous task, if not impossible.

METHODS

Datasets

The two independent datasets used here were gathered from studies published by researchers at the Hamilton Glaucoma Center of the University of California at San Diego⁶ (the Hamilton dataset) and by researchers at the Department of Ophthalmology of the University of Pittsburgh School of Medicine¹⁴ (the Pittsburgh dataset). Both reports stated that the studies followed the principles of the Declaration of Helsinki and were in compliance with the Health Insurance Portability and Accountability Act.

The Hamilton dataset⁶ consists of measurements of neuro-retinal rim area in the six sectors of the optic disc and corresponding visual sensitivities expressed in linear (1/Lambert) units.² Data for the superior-temporal sector of the optic disc (45°–90°) were selected to allow direct comparisons with the original work (see Fig. 2 in the study of Racette et al.⁶). Measurements were taken for 91 healthy eyes, 77 eyes at higher risk for development of glaucoma due to ocular hypertension, 125 eyes with suspected glaucoma, and 92 eyes with definite glaucoma. All subjects were part of the Diagnostic Innovations in Glaucoma Study (DIGS).

The Pittsburgh dataset¹⁴ consists of measurements of average RNFL thickness and relative visual field sensitivity.⁷ Measurements were taken for 72 eyes free of disease and 40 eyes with glaucoma. The data points were digitized from the graph in the top panel of Figure 1 from the report of Wollstein et al.¹⁴ Six points corresponding to eyes free of disease could not be digitized, due to overlapping, so the sample size of eyes free of disease used here was 66.

Statistical Methods

Racette et al.⁶ followed the approach of Garway-Heath et al.,² and used linear and polynomial regression to ask whether a linear or a decibel (dB) scale for visual sensitivities described better its relationship with rim area. Wollstein et al. asked a different question and used segmented regression to find the RNFL thickness value below which visual field damage should be detectable.¹⁴ Given their different research questions, it is

reasonable that they used different statistical methods. Equally here, different methods were applied to each dataset.

The Linearity Analysis. Three linear fits were obtained for the Hamilton data. The first was a linear OLS fit (simple linear regression) of visual sensitivity on rim area, as in the original study.⁶ The second was a linear OLS fit of rim area on visual sensitivity, as if the axes were swapped. The third was a linear SMA fit, which accounts for the fact that rim area and visual sensitivity are subject to measurement errors and intersubject differences. The key assumption in the linear SMA fit is that the ratio from the variance of measurement errors and intersubject differences in structure to that in function equals the ratio between variances of the observed structural and functional measures. In other words, SMA assumes that the signal-to-noise ratios in structure and function are equal. A review of OLS, SMA, and other regression methods, and their fundamental differences, can be found elsewhere.¹¹ The result for SMA does not change when the axes are swapped.

The Broken-Line Analysis. Four broken-line fits were obtained for the Pittsburgh data. The first was a broken-line fit of relative visual field sensitivity on RNFL thickness, as in the original study.¹⁴ The second was a broken-line fit of RNFL thickness on relative visual field sensitivity, as if the axes were swapped. The third and fourth were as the first and second fits, but after a change in the units of measure. Sensitivity was converted first to linear units,⁷ and then RNFL thickness and sensitivity values were divided by the average over the 66 eyes free of disease and multiplied by 100 to get percent of mean normal.

The Nonparametric Analysis. For the illustrative purpose of this study, a nonparametric local-linear fit, known as loess,¹³ also was obtained for both datasets after transforming all data to percent normal in the linear scale. The loess fit consists of estimating at each possible value of the explanatory variable the median value of the response by performing weighted least-squares linear regression locally. Locally means that only a portion of the range of possible values around the estimation point is used. The loess fit is determined by the specification of the smoothing parameter that determines which points are to be used in each local linear fit and their corresponding weights. Its value in each fit was selected manually following the guidelines of Cleveland.¹⁵

Even though loess fitting is a nonparametric procedure, it does not escape the need for assumptions, which are the same as for linear OLS regression. That is, all measurement errors and intersubject differences are attributed to the response variable. For this reason, two fits were obtained for each dataset: one of function on structure and one of structure on function. As association between structure and function was suspected to change from glaucomatous to healthy eyes, additional loess fits were obtained separately for eyes free of disease and eyes with glaucoma from both studies. For the Hamilton dataset, the 202 suspect eyes were removed, as this group is likely to contain healthy and glaucomatous eyes. As a control check on the manual selection of the smoothing parameter, loess fits also were obtained using a cross-validation method^{16,17} for automatic selection of the smoothing parameter. The fits obtained for automatic selection (not shown here) were less smooth than for manual selection, but the same overall patterns were found.

RESULTS

The Linearity Analysis

Figure 1 shows different fits to the Hamilton dataset.⁶ The three linear fits agree qualitatively in that visual sensitivity

FIGURE 1. Apparent linear association between structure and function. (a–d) show the same data points, visual sensitivity in linear scale (1/Lambert) and rim area in mm². (a) shows the linear OLS fit of sensitivity on rim area (compare with upper right panel of Fig. 2 in study of Racette et al.⁶). (b) shows the same as in (a), but after the axes were swapped. (c) shows, in addition to the fit in (b), the linear OLS fit of rim area on sensitivity. (d) shows the two linear OLS fits in (c) and the SMA linear fit. The same SMA linear fit would be obtained if the axes were swapped back.

increases, on average, with rim area. Nevertheless, the estimated slopes are very different from each other. The slope for the linear OLS fit of sensitivity on rim area (Figs. 1a–d, solid lines) was 360 μm^2 per 1/Lambert (significantly different from 0 with $P < 0.0001$).

The slope for the linear OLS fit of rim area on sensitivity obtained after the axes were swapped (Figs. 1c, 1d, dashed lines) was 34 μm^2 per 1/Lambert ($P < 0.0001$); 10 times smaller. The slope for the SMA fit (Fig. 1d, dotted line) was 110 μm^2 per 1/Lambert. The slope estimated with SMA fit was found to be significantly different from slopes estimated with OLS fit of sensitivity on rim area ($P < 0.0001$) and of rim area on sensitivity ($P < 0.0001$), with the “common slopes test” for SMA.¹⁸ Therefore, from the same dataset, three statistically

different apparent associations were obtained from three different methods.

The large difference between estimated slopes is due to the fact that data are very poorly correlated, with $R^2 = 0.09$; well below the conventional cutoff for strong correlation, $R^2 = 0.50$. With such low correlation, when the linear model explains less than half the variation in the data, predictions for a particular instance are likely to be severely biased. The three fits agree in the qualitative assessment that, on average, rim area increases with visual sensitivity. Therefore, structure and function, indeed, seem to be associated, even if the association is very weak. If there were no association, then the estimated slope in the OLS fits would not be significantly different from zero. Thus, the solid line in Figure 1d would be horizontal and the

FIGURE 2. Finding the breakpoint. (a–d) show the same data points, relative visual field sensitivity and RNFL thickness. In (a–c) the units were in μm for RNFL thickness and dB for sensitivity. (a) shows a broken-line fit of sensitivity on thickness (compare with upper panel of Fig. 1 in the study of Wollstein et al.¹⁴). (b) shows the same data as in (a), but after the axes were swapped. (c) shows, in addition to the fit in (b), the broken-line fit of thickness on sensitivity. (d) is a replication of (c), but after changing units to percent normal (see Methods).

dashed line would be vertical. The SMA fitted line in Figure 1d is between the two OLS fits. By construction, and in contrast to OLS fits, slopes estimated with SMA are systematically different from zero even when there is no association. Therefore, SMA is not an appropriate method for assessing associations between structure and function.

The Broken-Line Analysis

Figure 2 shows different broken-line fits to the Pittsburgh dataset.¹⁴ Six data points were lost in the digitization process. Details for Figures 2a to 2c are the same as those in Figure 1, but for a different dataset and statistical model (the broken-line model as opposed to the linear model). Thus, the solid lines in

Figures 2a to 2c represent the broken-line fit of sensitivity on thickness, and the dashed line in Figure 2c represents the broken-line fit of thickness on sensitivity obtained after the axes were swapped. In Figure 2d the broken-line fits for thickness on sensitivity (dashed line) and sensitivity on thickness (solid line) were calculated after data were converted to percent of mean normal (see Methods for the broken-line analysis for details).

The impact of differences in broken-line fits on statistical inferences was explored by answering three questions asked in the form of statistical hypotheses (with a significance level fixed at 0.05), as in the original study.¹⁴ The first question was whether the slope before the breakpoint was equal to zero. The second question was whether the slope after was equal to

TABLE. Finding the Breakpoint With Broken-Line Fits at Four Different Statistical Settings

	Breakpoint in μm	Slope Before = 0?	Slope After = 0?	Before = After?
Sensitivity on thickness*	75	No	Yes	No
Thickness on sensitivity	61	Yes	No	Yes
Sensitivity on thickness in % normal†	77	No	Yes	Yes
Thickness on sensitivity in % normal	81	No	No	Yes

For each setting, the corresponding thickness value at the estimated breakpoint is shown in μm along with the results, at a significance level of 0.05, of three statistical tests of hypotheses. The mean normal thickness was estimated at 90.5 μm from the 66 eyes free of disease. The first hypothesis was that the slope before the breakpoint equals zero, the second was that the slope after the breakpoint equals zero, and the third was that the slopes before and after the breakpoint are equal. The first two rows show results for relative field sensitivity in dB and RNFL thickness in μm (as in Figs. 2a–c). The third and fourth rows are the same as first and second, respectively, but after a change in the units of measure to percent normal (as in Fig. 2d).

* This analysis is as the original¹⁴ except for six points that are missing here after digitation.

† This analysis also is as the original secondary analysis¹⁴ except for the six missing points and the scaling constant of 100/90.5 μm used to convert to percent normal. The Davies' test here gave a different result than the original analysis¹⁴ as a result of the six missing points.

zero. The third question was whether the slopes before and after the breakpoint were equal to each other. For the last hypothesis the Davies' test¹⁹ was used. The Table shows the RNFL thickness for the breakpoint obtained from each broken-line fit in Figure 2, along with the corresponding answers to each of the three questions about slopes.

Four different analyses led to four different conclusions. Two of the analyses (rows 2 and 3 in the Table) were inconsistent. The other two (rows 1 and 4 in the Table) contradicted each other: the first was consistent with a breakdown in the linearity of the association,¹⁴ the fourth was consistent with a linear association between structure and function.²

The Nonparametric Analysis

As an alternative way to explore apparent associations between structure and function, loess¹³ fits were used to examine visually whether or not there was an apparent change in the association between observations of structural and functional measures. Figure 3 shows loess fits for the Hamilton and Pittsburgh datasets.

The range in the values on the axes of Figure 3b have been changed from those in Figure 2d to facilitate visual comparisons with the other graphs in Figures 3 and 4. There seemed to be a change in the association between the structural and functional measures. This change was suspected to be due to a difference between healthy and glaucomatous eyes, and, therefore, loess fits were obtained separately for each group in each dataset.

Figure 4 shows loess fits for glaucomatous eyes (left panels) and healthy eyes (right panels). Upper panels show fits for the Hamilton dataset and lower panels for the Pittsburgh dataset. The data in the upper panels are the same as in Figure 3a, but split into glaucomatous eyes (Fig. 4a) and healthy eyes (Fig. 4b). Equally, the data in the lower panels are the same as in Figure 3b, but split into glaucomatous eyes (Fig. 4c) and healthy eyes (Fig. 4d).

For healthy eyes (Figs. 4b, 4d), all loess fits were roughly constant. These fits are to be expected if structure and function were independent of each other. The RNFL thickness of glaucomatous eyes increased, on average, with relative visual field sensitivity (see Fig. 4c). The association between rim area and sensitivity, however, was unclear: according to loess fit of sensitivity on rim area, sensitivity increased, on

FIGURE 3. Apparent nonlinear association between structure and function. (a) shows data points from the Hamilton dataset, and corresponding loess fits to definitely glaucomatous eyes and healthy eyes for rim area and sensitivity. (b) shows data points from the Pittsburgh dataset, and the loess fits to glaucomatous eyes and healthy eyes for RNFL thickness and relative visual field sensitivity.

FIGURE 4. Apparent association between structure and function in healthy and glaucomatous eyes. (a, b) show data points and the corresponding loess fits to eyes with definite glaucoma and healthy eyes, respectively, from the Hamilton dataset. The data in (a, b) are the same as for Figure 3a. (c, d) show data points and the corresponding loess fits to data from eyes with glaucoma and healthy eyes, respectively, from the Pittsburgh dataset. The data in (c, d) are the same as for Figure 3b.

average, until rim area was 100% of mean normal when it started to decrease. This is an unexpected result that points out a limitation of loess fits, or any other local fit, when data points are scarce at extreme values.

DISCUSSION

Statistical analyses based on different fitting methods were compared against each other with two different datasets. Results differed across fitting methods, depending on the analysis setup and fitting procedure. The intention of this study was to get a better understanding of the limitations of the statistical models used in studies of association between structure and function; in particular, about what can be inferred and what cannot from them, and whether one can be sure about such inferences. Although pitfalls and limitations of widely used regression methods have been studied repeatedly in medical-statistics literature, and other areas, such as chemistry⁹ and biology,^{11,18} they often are not recognized in studies of association between structure and function in glaucoma.

Linear Regression Methods

Linear fits obtained with different procedures can be very different from each other (see Fig. 1d). The main reason for such inconsistencies is that the large amount of measurement errors and intersubject differences interact differently with the unmet assumptions of each fitting procedure. Very different fits also would be obtained if the same fitting procedure was applied to different datasets, partially because of measurement errors and intersubject differences, and partially because the true association may not be linear (see Fig. 4, where apparent association seems to be different for healthy eyes than for glaucomatous eyes). If structure and function are independent from each other for healthy eyes, but positively associated for glaucomatous eyes, then linear fits to datasets containing larger proportions of controls will tend to yield smaller slope estimates.

Of all fits considered here, only SMA does account for the fact that there is no error-free explanatory variable, and both structure and function have measurement errors and intersubject differences. Nevertheless, SMA assumes a specific way in

which such errors relate to each other. There is no strong reason to believe that the SMA assumption is any better than the OLS assumption. Other methods, such as the Deming regression²⁰ (or a weighted version), may be more appropriate if the nature of the relationship between measurement errors and intersubject differences in structure and function becomes better understood.

There are many more linear-regression methods. In a recent review, Ludbrook accounted for more than 25 variations.¹⁰ Among others, Passing-Bablok rank-order regression^{21–23} has been used widely in the clinical domain. Unlike the methods discussed here, Passing-Bablok regression does not require specific assumptions regarding the distribution of measurement error. Furthermore, the estimation of the line parameters is robust to outliers.¹⁰ Nevertheless, since the Gaussian-distribution assumptions made by classic regression methods have been shown to have little effect on the estimation of parameters,²⁴ it does not seem to be always advantageous to use Passing-Bablok regression over, for example, SMA or Deming regression.²⁰ Readers are kindly referred to the reports by Ripley and Thompson,⁹ and other investigators^{10,11,20,24,25} in which linear fitting procedures for bivariate data and their potential pitfalls are presented in a more detailed and systematic manner. Yet, the problem remains that the true association may not be well described by a linear model.

Nonparametric Methods

Of all fits considered here, only loess does not assume a particular shape of the data, and it can give insights as to whether the association truly is linear or not (see Figs. 3, 4). Sometimes, however, it can give unexpected, nonsensical fits, depending on the dataset (Fig. 4a, solid line). Loess is a method known to be unreliable at extreme values when data are highly variable. The fits in Figures 4b and 4d, however, are consistent with the assumption that structure and function are independent of each other for eyes free of disease.

This study illustrates that no fitting method is intrinsically superior to any other. Given the magnitudes of measurement errors and intersubject differences in the data, it is impossible to be sure that any model is an accurate description of the true underlying association between structure and function. An assessment as such would not be possible to make until errors in measurements are greatly reduced, and biological variability in the data are properly accounted for. These observations extend previous ones by Racette et al. on inconsistencies between statistical analyses of apparent association between structure and function.⁶

Units of Measure

In addition to the statistical fitting procedure, the units of measure used have an impact on the apparent association between structure and function. An example was given here by replicating, after changes in the units of measure to percent normal, the broken-line analysis to the Pittsburgh dataset (see Table). Four analyses gave four different answers, two that were inconsistent and two that contradicted each other. There is no strong reason to believe one analysis is any more valid than the other.

Complex Statistical Models

Alternatives to linear models that use one global index for structure and one global index for function have been developed. Latent class regression models (Bilonick, et al. *IOVS* 2012;53:ARVO E-Abstract 4629), for instance, have been used to extract the “latent information” shared among

different indices for glaucoma (Wollstein, et al. *IOVS* 2012;53:ARVO E-Abstract 218). This approach assumes that there is an unknown, unobservable variable that quantifies the “true” stage of the disease and that can be reconstructed from the observation of the imperfect indices for glaucoma. Other approaches have used radial basis functions under a Bayesian framework,^{26,27} multiple linear regression based on principal component analysis²⁸ to estimate sensitivity from structural measures at each location of the visual field, and multilayer artificial neural networks²⁹ to integrate structure and function information also at each location of the visual field. These methods, however promising, are subject to other challenges, such as overfitting,²⁶ which are yet to be assessed.

CONCLUSION

Because of the magnitudes of measurement errors and intersubject differences, ambiguous inferences can be drawn accidentally from studies of the association between structure and function. In turn, this can lead to inappropriate study designs or recommendations for clinical practice. For instance, the arguable assumption that structural abnormality always occurs in glaucomatous eyes before visual function declines could lead to the recommendation of using imaging devices for detection of early glaucoma without testing for visual function. Given the large biological differences between subjects, and regardless of whether structural damage does occur before functional damage, if a recommendation as such were followed, then a substantial percent of subjects with glaucoma would pass undetected.³⁰ Awareness of the limitations of any statistical method is necessary to avoid being misled by spurious apparent associations between structure and function in glaucoma.

Acknowledgments

The authors thank David Foster for useful discussion on loess analysis and, along with Lyne Racette and Paul Artes, for critically reading a draft of the manuscript.

Supported by NIH Grant R01EY007716 (WHS), and by the National Institute for Health Research (NIHR) Biomedical Research Centre based at Moorfields Eye Hospital NHS Foundation Trust and UCL Institute of Ophthalmology (RM). Linda Zangwill, Felipe Medeiros, and Pamela Sample provided the data from the Diagnostic Innovations in Glaucoma Study (DIGS), supported by NEI Grants U10EY14267, EY019869, EY08208, EY11008, and EY13959. The authors alone are responsible for the content and writing of the paper.

Disclosure: **I. Marín-Franch**, None; **R. Malik**, None; **D.P. Crabb**, None; **W.H. Swanson**, None

References

1. Garway-Heath DE, Poinoosawmy D, Fitzke FW, Hitchings RA. Mapping the visual field to the optic disc in normal tension glaucoma eyes. *Amer Acad Ophthalmol*. 2000;107:1809–1815.
2. Garway-Heath DE, Holder GE, Fitzke FW, Hitchings RA. Relationship between electrophysiological, psychophysical, and anatomical measurements in glaucoma. *Invest Ophthalmol Vis Sci*. 2002;43:2213–2220.
3. Jansonius NM, Schiefer J, Nevalainen J, Paetzold J, Schiefer U. A mathematical model for describing the retinal nerve fiber bundle trajectories in the human eye: average course, variability, and influence of refraction, optic disc size and optic disc position. *Exp Eye Res*. 2012;105:70–78.

4. Airaksinen PJ, Drance SM, Douglas GR, Schulzer M, Wijsman K. Visual field and retinal nerve fiber layer comparisons in glaucoma. *Arch Ophthalmol*. 1985;103:205-207.
5. Harwerth RS, Vilupuru AS, Rangaswamy NV, Smith EL III. The relationship between nerve fiber layer and perimetry measurements. *Invest Ophthalmol Vis Sci*. 2007;48:763-773.
6. Racette L, Medeiros FA, Bowd C, Zangwill LM, Weinreb RN, Sample PA. The impact of the perimetric measurement scale, sample composition, and statistical method on the structure-function relationship in glaucoma. *J Glaucoma*. 2007;16:676-684.
7. Hood DC, Kardon RH. A framework for comparing structural and functional measures of glaucomatous damage. *Prog Retin Eye Res*. 2007;26:688-710.
8. Hood DC, Anderson SC, Wall M, Raza AS, Kardon RH. A test of a linear model of glaucomatous structure-function loss reveals sources of variability in retinal nerve fiber and visual field measurements. *Invest Ophthalmol Vis Sci*. 2009;50:4254-4266.
9. Ripley BD, Thompson M. Regression techniques for the detection of analytical bias. *Analyst*. 1987;112:377-383.
10. Ludbrook J. Linear regression analysis for comparing two measurers or methods of measurement: but which regression? *Clin Exp Pharmacol Physiol*. 2010;37:692-699.
11. Warton DI, Wright IJ, Falster DS, Westoby M. Bivariate line fitting methods for allometry. *Biol Rev Camb Philos Soc*. 2006;81:259-291.
12. Muggeo VMR. Estimating regression models with unknown break-points. *Stat Med*. 2003;22:3055-3071.
13. Cleveland WS. Robust locally weighted regression and smoothing scatterplots. *J Amer Stat Assoc*. 1979;74:829-836.
14. Wollstein G, Kagemann L, Bilonick RA, et al. Retinal nerve fiber layer and visual function loss in glaucoma: the tipping point. *Brit J Ophthalmol*. 2011;96:47-52.
15. Cleveland WS. *The Elements of Graphing Data*. Summit, NJ: Hobart Press; 1985: chap 3.
16. Loader CR. Bandwidth selection: classical or plug-in? *Ann Statist*. 1999;27:415-438.
17. Xia Y, Li WK. Asymptotic behavior of bandwidth selected by the cross-validation method for local polynomial fitting. *J Multiv Anal*. 2002;83:265-287.
18. Warton DI, Weber NC. Common slope tests for bivariate errors-in-variables models. *Biometric J*. 2002;44:161-174.
19. Davies RB. Hypothesis testing when a nuisance parameter is present only under the alternative. *Biometrika*. 1987;74:33-43.
20. Linnet K. Evaluation of regression procedures for methods comparison studies. *Clin Chem*. 1993;39:424-432.
21. Passing H, Bablok WA. New biometrical procedure for testing the equality of measurements from two different analytical methods. Application of linear regression procedures for method comparison studies in clinical chemistry, part I. *J Clin Chem Clin Biochem*. 1983;21:709-720.
22. Passing H, Bablok W. Comparison of several regression procedures for method comparison studies and determination of sample sizes. Application of linear regression procedures for method comparison studies in clinical chemistry, part II. *J Clin Chem Clin Biochem*. 1984;22:431-445.
23. Passing H, Bablok W, Bender R, Schneider BA. General regression procedure for method transformation. Application of linear regression procedures for method comparison studies in clinical chemistry, part III. *J Clin Chem Clin Biochem*. 1988;26:783-790.
24. Knight JL. Nonnormal errors and the distribution of OLS and 2SLS structural estimators. *Econ Theory*. 1986;2:75-106.
25. Martin RE. General Deming regression for estimating systematic bias and its confidence interval in method-comparison studies. *Clin Chem*. 2000;46:100-104.
26. Zhu H, Crabb DP, Schlottmann PG, et al. Predicting visual function from the measurements of retinal nerve fiber layer structure. *Invest Ophthalmol Vis Sci*. 2010;51:5657-5666.
27. Zhu H, Crabb DP, Fredette MJ, Anderson DR, Garway-Heath DE. Quantifying discordance between structure and function measurements in the clinical assessment of glaucoma. *Arch Ophthalmol*. 2011;129:1167-1174.
28. Zhang X, Bregman CJ, Raza AS, De Moraes G, Hood DC. Deriving visual field loss based upon OCT of inner retinal thicknesses of the macula. *Biomed Opt Express*. 2011;2:1734-1742.
29. Bizios D, Heijl A, Bengtsson B. Integration and fusion of standard automated perimetry and optical coherence tomography data for improved automated glaucoma diagnostics. *BMC Ophthalmology*. 2011;11:20.
30. Malik R, Swanson WH, Garway-Heath DE. 'Structure-function relationship' in glaucoma: past thinking and current concepts. *Clin Exp Ophthalmol*. 2012;40:369-380.