ONLINE FIRST

Quantifying Discordance Between Structure and Function Measurements in the Clinical Assessment of Glaucoma

Haogang Zhu, PhD; David P. Crabb, PhD; Marie-Josée Fredette, MD; Douglas R. Anderson, MD; David F. Garway-Heath, MD

Objective: To evaluate a new method of quantifying and visualizing discordance between structural and functional measurements in glaucomatous eyes by predicting the visual field (VF) from retinal nerve fiber layer thickness (RNFLT) using a bayesian radial basis function.

Methods: Five GDx VCC RNFLT scans and 5 Humphrey 24-2 Swedish Interactive Thresholding Algorithm VF tests were performed for 50 glaucomatous eyes from 50 patients. A best-available estimate (BAE) of the true VF was calculated as the pointwise median of these 5 replications. This BAE VF was compared with every RNFLT-predicted VF from the bayesian radial basis function and every measured VF. Predictability of VFs from RNFLT was established from previous data. A structure-function pattern discordance map and a structure-function discordance index (scores of 0-1) were established from the predictability limits for each structure-function measurement pair to quantify and visualize the discordance between the structure-predicted and measured VFs.

Results: The mean absolute difference between the structure-predicted and BAE VFs was 3.9 dB. The mean absolute difference between measured and BAE VFs was 2.6 dB. The mean (SD) structure-function discordance index score was 0.34 (0.11). Ninety-seven (39%) of the structure-predicted VFs showed significant discordance (structure-function discordance index score >0.3) from measured VFs.

Conclusions: On average, the bayesian radial basis function predicts the BAE VF from RNFLT slightly less well than a measured VF from the 5 VFs composing the BAE VF. The pattern discordance map highlights locations with structure-function discordance, with the structure-function discordance index providing a summary index. These tools may help clinicians trust the mutually confirmatory structure-function measurements with good concordance or identify unreliable ones with poor concordance.

Arch Ophthalmol. 2011;129(9):1167-1174.
Published online May 9, 2011.
doi:10.1001/archophthalmol.2011.112

Author Affiliations are listed at the end of this article.

GLAUCOMA IS A PROGRESSIVE optic neuropathy in which structural damage is evident at the optic nerve head and the retinal nerve fiber layer (RNFL). This structural damage results in the loss of visual function during the progression of the disease. Glaucoma is clinically evaluated by various tests that, in general, examine structural or functional properties of the optic disc or retina. Ideally, the functional loss would be precisely predictable by the magnitude and configuration of the structural loss.[1] However, the relationship between the structural and functional measurements is still not well understood, mainly owing to the poor precision and accuracy of the current clinical devices measuring the structural and functional deficits. Particularly, various structural measures use optical imaging techniques to examine optic nerve head parameters or RNFL thickness (RNFLT) as surrogates for the biological variable of interest, namely the number of (functioning) retinal ganglion cells. Visual function can be examined with visual field (VF) tests of various kinds, some of which are similar enough to be grouped under the term *standard automated perimetry* (SAP). Standard automated perimetry is the clinical cornerstone in the assessment of glaucoma but is also subject to considerable measurement variability and inaccuracy.[2] Despite their limitations, these techniques are presently the state-of-the-art tools for the diagnosis and management of glaucoma.

Downloaded from www.archophthalmol.com at , on December 27, 2011
©2011 American Medical Association. All rights reserved.

When multiple types of measurement assessing glaucoma damage are available for the same patient, an important question is whether they are consistent with each other; if not, to what degree and why are they discordant? Thinking about this would help clinicians to avoid being misled by inaccurate or imprecise measurements. Recently, clinical software has been developed that presents structure and function classification analysis (probability levels) for corresponding regions of the VF and optic nerve head (Heidelberg Eye Explorer version 1.7.0; Heidelberg Engineering, Heidelberg, Germany). This enables visualization of structure-function agreement, or spatial concordance, at an optic nerve head sector level in both the structural and functional measurements, but the degree of concordance is not quantified.

Attempts have been made to correlate regions from the structural measures (eg, mean RNFLT in 6 sectors) and groups of or individual points in the VF and to assess the curve-linear (eg, log-linear) or monotonic association between the 2 variables via R^2, Pearson, or Spearman coefficients.[3-13] As methods of examining and demonstrating regional structure-function relationships, these approaches make restricted assumptions and by their nature do not quantify the frequency or magnitude of structure-function agreement (concordance) or disagreement (discordance) at a high spatial resolution (for example, at each VF test location). First, the analyses suffer from considerable data reduction by summarizing the structural image into a few discrete measurements (for instance, the sectoral RNFLT derived from the GDx VCC [Carl Zeiss Meditec, Dublin, California]) or a grouping of VF locations, so the spatial resolution of the measurements is lost. Second, these analyses assume and are limited to a particular shape of the structure-function relationship, while in reality this relationship may be more complex with the nature of the association possibly changing across the stages of the disease. Third, these studies treat individual points from structure and function measurements as independent and fail to address the inherent spatial relationships within the individual measurements. Last but not least, owing to measurement variability and anatomical variation, the structure-function relationship in these studies is reported as a statistical association based on measurements from a population, which only approximates the relationship of any particular individual in the population. However, the structure-function discordance should be quantified for the measurements of individual patients and the method should be generalized to the individuals not included in the data set used to establish the structure-function relationship.

In contrast to association-based methods, a more flexible nonlinear prediction model, the bayesian radial basis function (BRBF), has been developed and validated to predict the VF from the RNFLT.[1] The model was built and tested on the measurements of 535 eyes from 535 subjects from 3 independent centers. It was shown that, on average, the BRBF model can predict a patient's VF from the RNFLT almost as well as another VF from the patient. Moreover, the BRBF derived a high-resolution structure-function relationship that is consistent with the known typical location of retinal ganglion cells that project their axons toward particular meridians of the optic disc.[14] The VF predicted from the RNFLT by the BRBF can be understood as the expectation or representation of retinal structure in the domain of the VF. Therefore, structure-function discordance is converted to discordance between the structure-expected VF and the measured VF.

The purpose of this study was to use the BRBF method to predict the expected result of a VF test based on structural measurements. Thereafter, an analysis was undertaken to quantify and visualize the discordance between an individual VF test and the expectation based on structural information. This method is demonstrated, on a test-retest data set, in hopes of providing a basis for a clinical tool that gives an alert for inconsistency of a VF test with structural measurements.

METHODS

SUBJECTS

A test-retest data set was used in this study. Fifty patients were recruited from a hospital-based glaucoma ophthalmology practice (Bascom Palmer Eye Institute, Anne Bates Leach Eye Hospital, University of Miami Miller School of Medicine, Miami, Florida). Glaucoma was diagnosed by detailed medical and ocular histories and by standard ophthalmic examination. The ophthalmic examination included best-corrected visual acuity, anterior segment examination, Goldmann applanation tonometry, dilated fundus examination, and SAP with the 24-2 Swedish Interactive Thresholding Algorithm standard program of the Humphrey field analyzer (Carl Zeiss Meditec). The data set covers a wide spectrum of glaucoma severity from early to advanced stage, judged by the patients' treating physicians to be under control with stability of nerve damage and VF loss. Inclusion criteria for these patients included the following: (1) aged 18 years or older; (2) controlled intraocular pressures; (3) best-corrected visual acuity of 20/40 or better; (4) less than 5 diopters of spherical refractive error and 3 diopters of cylindrical refractive error; (5) pupil diameter of 2 mm or more; (6) no history of ocular or neurological disease or surgery that might produce nonglaucomatous structural or functional abnormality; (7) no history of amblyopia; (8) mental and physical capacity to perform the tests; and (9) willingness to participate in the study. Only 1 eye had to fit the eligibility criteria for the patient to enter the study. If both eyes met the eligibility criteria, the study eye was selected to ensure representation of a wide spectrum of the disease (an even distribution of selected eyes using the VF mean deviation).

Each subject attended 5 sessions within an 8-week period. During each of the 5 sessions, the subjects had SAP, scanning laser polarimetry, and optical coherence tomography (OCT) measurements done. The SAP was performed with the 24-2 Swedish Interactive Thresholding Algorithm standard program of the Humphrey field analyzer with the Goldmann size III stimulus. The SAP test used to establish the diagnosis and eligibility was not included in the 5 repeated tests obtained for the study experiments. The scanning laser polarimetry measurements were taken with the GDx VCC. All scans were acquired through undilated pupils. In this data set, measurement reliability indices were similar to those of the data set used to establish the BRBF model[1]: for VF tests, fixation losses of 20% or less, a false-positive rate of 20% or less, and a false-negative rate of 33% or less; for GDx VCC images, an image quality score of 8 or higher and a typical scan score of 80 or higher. The OCT images were taken with the Stratus OCT (Carl Zeiss Meditec) using the 3.4-mm fast-scan protocol.

Downloaded from www.archophthalmol.com at , on December 27, 2011
©2011 American Medical Association. All rights reserved.

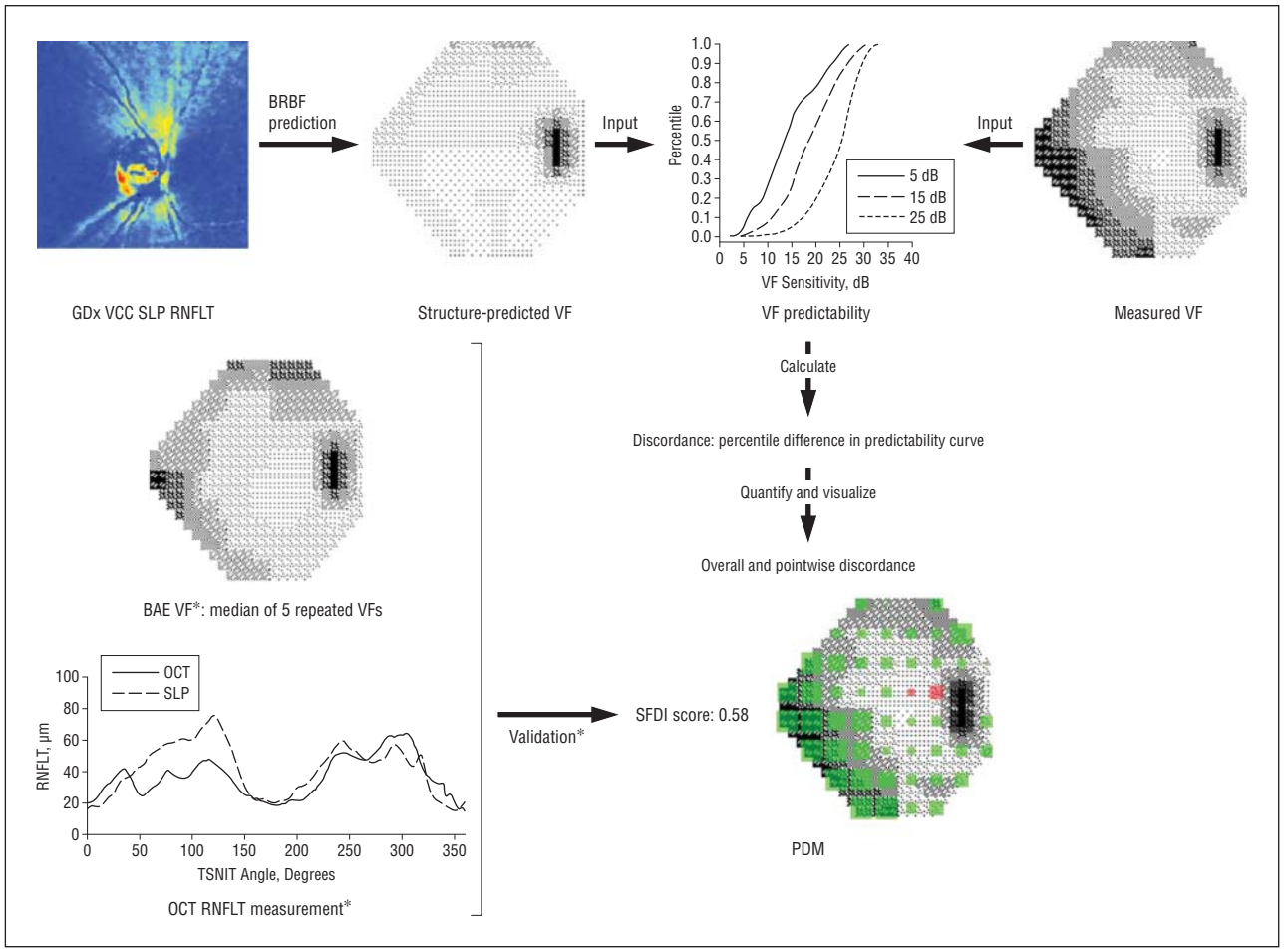


Figure 1. Flowchart summarizing the analysis in the study. *The data in the validation step are only for the purpose of validating the method and are not necessary when using the proposed method in practice. BAE indicates best-available estimate; BRBF, Bayesian radial basis function; OCT, optical coherence tomography; PDM, pattern discordance map; RNFLT, retinal nerve fiber layer thickness; SFDI, structure-function discordance index; SLP, scanning laser polarimetry; TSNIT, temporal-superior-nasal-inferior-temporal; and VF, visual field.

The study followed the tenets of the Declaration of Helsinki, and the protocol was approved by the Human Subjects Institutional Review Board of the University of Miami. Patients agreed to participate as subjects in the study and signed an informed consent form after explanation of the nature and possible consequences of the study. The data were collected in the last half of 2004. They were later anonymized and transferred to a secure database held at City University London, London, England, for the conduct of the analysis.

ANALYSIS

The analysis has been summarized as a flowchart in **Figure 1**. The details are given in the following sections.

Structure-Predicted VF

The structure-function relationship was previously modeled with a BRBF, which was used to predict the VF from the RNFLT profile from the GDx VCC. The model has been trained on RNFLT and VF measurements of 229 eyes, and it derived a structure-function relationship consistent with the retinal anatomy such that the structure-predicted VF (Figure 1) can be understood as the RNFLT representation in the VF domain.[1] The structure-function discordance, as described later, is defined as the difference between this structure-predicted VF and the measured VF (Figure 1).

Best-Available Estimate of True VF

To compare the predicted and measured VFs, the true VF should be known; however, high variability of testing means that the true VF cannot be known from any particular VF test. Therefore, a best-available estimate (BAE) (Figure 1) of the true VF was calculated as the pointwise median of the 5 repeated VFs. It was then assumed that the BAE VF provides a closer estimate of the true VF than any single measurement. The BAE VF was used as the reference to identify the location, extent, and possible cause of any discordance when the structure-predicted VF and the measured VF were discordant. However, the method proposed in this study is not constrained by the availability of a BAE VF.

The difference between each structure-predicted VF and the BAE VF was quantified. This difference was compared with the difference between each VF from the test-retest set and the BAE VF.

VF Predictability and Quantification of Structure-Function Discordance

The discordance between the structure-predicted VF and the measured VF may be easily calculated as the pointwise difference of the 2 sensitivities. However, the predictability of the VF varies with the VF sensitivity itself. In **Figure 2**A, as re-

Downloaded from www.archophthalmol.com at , on December 27, 2011
©2011 American Medical Association. All rights reserved.

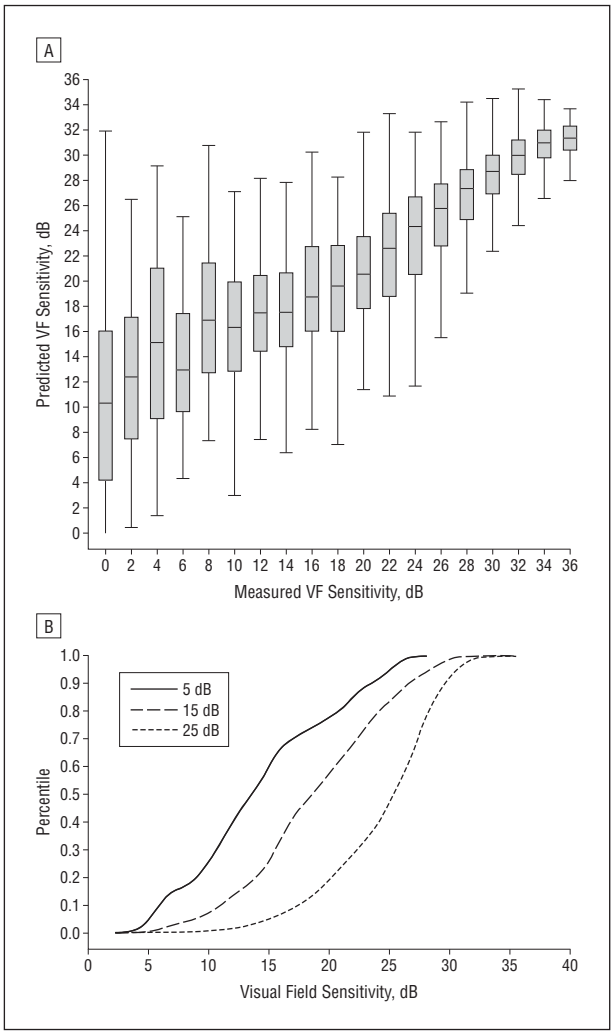


Figure 2. Bayesian radial basis function predictability and percentile curves. A, Bayesian radial basis function prediction (predictability) ranges at different visual field (VF) sensitivities. Error bars indicate 90% (5%-95%) prediction limits; top and bottom of boxes, interquartile (25%-75%) prediction limits. B, Percentile curves of predictability at measured sensitivities of 5, 15, and 25 dB.

ported previously,[1] the range of prediction errors decreases with a higher measured VF sensitivity. This means that a higher VF sensitivity is more predictable than a lower sensitivity; this needs to be taken into account in the calculation of the discordance. The VF predictability (Figure 1) can be quantified as the percentile of the predicted VF sensitivity for any given measured VF sensitivity. This maps the VF sensitivity in decibels to a prediction percentile ranging from 0 to 1. The percentile curves for the predicted VF sensitivity are illustrated in Figure 2B at the measured VF sensitivities of 5, 15, and 25 dB. The variable predictability at different measured sensitivities is illustrated by the different shapes of the percentile curves in Figure 2B: the corresponding curve shifts toward the right and has a steeper slope (increases from 0 to 1 more rapidly) at higher measured sensitivity. Note that the BRBF model, from which the predictability is derived, was previously developed with measurements selected with strict quality criteria.[1]

Due to the greater predictability at high sensitivity, any given difference between predicted and measured VF sensitivities has greater importance if the measured sensitivity is high than if it is low. To compensate for this, the discordance between the structure-predicted and measured VF sensitivities is calculated as the percentile difference on the predictability curve (corresponding to the measured sensitivity) between the structure-predicted and measured sensitivity values. For instance, for a 5-dB difference between predicted and measured VF sensitivities at a measured sensitivity of 5 dB, the discordance between a 0-dB prediction and 5-dB measurement is 5.1% (on the 5-dB predictability curve, Figure 2B), but for a 5-dB difference at a 15-dB measured sensitivity, the discordance between a 10-dB prediction and 15-dB measurement is significantly higher at 17.3% (on the 15-dB predictability curve, Figure 2B). Similarly, the discordance between a 20-dB prediction and 25-dB measurement is even higher: 30.1%. Therefore, by using the predictability percentiles, the difference between the structure-predicted and measured VF sensitivities in decibels is normalized with respect to the predictability at any given level of measured sensitivity.

The signed pointwise discordance (Figure 1) between the structure-predicted and measured VF sensitivities quantifies the structure-function discordance: the sign is positive if the RNFLT overestimates the measured VF and vice versa. The value ranges from –1 to +1.

Pattern Discordance Map and Structure-Function Discordance Index

The pointwise structure-function discordance can be visualized with a novel tool, the structure-function pattern discordance map (PDM) (Figure 1), with a technique similar to the Hinton diagram, which has been used to visualize the weight parameters in a multilayer perceptron neural network.[15] As shown in Figure 1, the structure-function PDM is a 24-2 grayscale representation of the VF with red or green squares superimposed on each location of the measured VF. A green square indicates that the structural measure did not predict as low a visual sensitivity as the actual test result (discordance >0), while a red square means that the structural measure predicted a greater visual defect than was found in VF testing (discordance <0). The size of the square represents the magnitude of the discordance. Like the pattern standard deviation map that is used in a Humphrey field analyzer VF chart to describe the deviation from the normative database, the PDM provides an easily understood clinical tool to identify the pointwise deviation of the structure-predicted VF from the measured VF.

To summarize the overall discordance, a structure-function discordance index (SFDI) (Figure 1) was defined as the mean of absolute (unsigned) pointwise discordance across all locations in the VF. It ranges between 0 and 1, where 0 indicates no discordance and 1 means complete discordance. The SFDI quantifies the average difference between the structure-predicted and measured VFs and therefore acts similar to a global index from a VF chart.

The SFDI and the PDM were used on the test-retest data set to evaluate the structure-function discordance in this glaucomatous sample. The BAE VF was used as the arbiter to judge which measurement is correct when the discordance was flagged by either of the tools. The OCT scans for the same eyes were also used to identify the possible source of discordance.

RESULTS

The mean absolute difference between the structure-predicted VF and the BAE VF was 3.9 dB (SD, 4.3 dB). This compares with the mean absolute difference between a single measured VF and the BAE VF of 2.6 dB (SD, 3.5 dB).

The SFDI score was calculated for each of the 5 pairs of structural and functional measurements taken in the

Downloaded from www.archophthalmol.com at , on December 27, 2011
©2011 American Medical Association. All rights reserved.

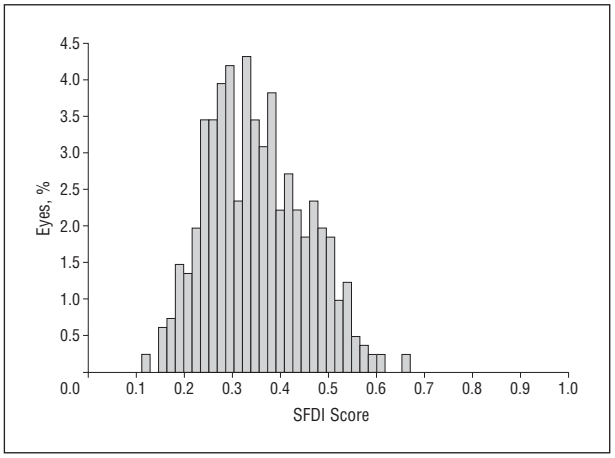


Figure 3. Histogram of structure-function discordance index (SFDI) scores in this test-retest sample of glaucomatous eyes.

same session. In this glaucomatous sample, the mean (SD) SFDI score was 0.34 (0.11), and the distribution of SFDI scores is shown in **Figure 3**. Although the SFDI score is a continuous variable with a suggestion of bimodal distribution, an arbitrary threshold (SFDI score of 0.30) was chosen to divide the measurements into 2 groups: a group with good overall concordance and the other with noticeable discordance. One hundred fifty-three measurements (61%) fall in the group with good concordance. The mean (SD) SFDI score was 0.25 (0.04) for the group of measurements with good overall structure-function concordance (SFDI score ≤0.30) compared with 0.41 (0.08) for the group with noticeable discordance (SFDI score >0.30).

Four examples with little, moderate, and significant structure-function discordance are illustrated in **Figures 4**, **5**, **6**, and **7**. The SFDI and structure-function PDM are demonstrated as clinical tools to identify the disagreement between structural and functional measurements in these 4 examples.

In Figure 4, the structure-predicted, measured, and BAE VFs all indicated an early defect in the nasal, nasal-superior, and nasal-inferior areas. The SFDI score of 0.15 demonstrates a good overall structure-function concordance, and the structure-function PDM showed no significant flags with large discordance squares.

The SFDI evaluates the overall structure-function discordance but may fail to pick up local discordance. For example, the SFDI score of 0.28 in Figure 5 is still within the arbitrary threshold of a satisfactory overall concordance, but the summary index did not capture the discordance in the nasal area that was flagged by the PDM. In this example, RNFLT did not predict the nasal defect in the measured VF; this disagreement was flagged by 2 large green discordance squares in the nasal area.

Figure 6 is an example with moderate overall discordance (SFDI score of 0.58), and the locations of the discordance were further described by the PDM. The RNFLT overestimated the measured VF especially in the inferior part of the VF and the single measured VF was more consistent with the BAE VF, showing that the structural measurement is more likely to be the cause of the discordance.

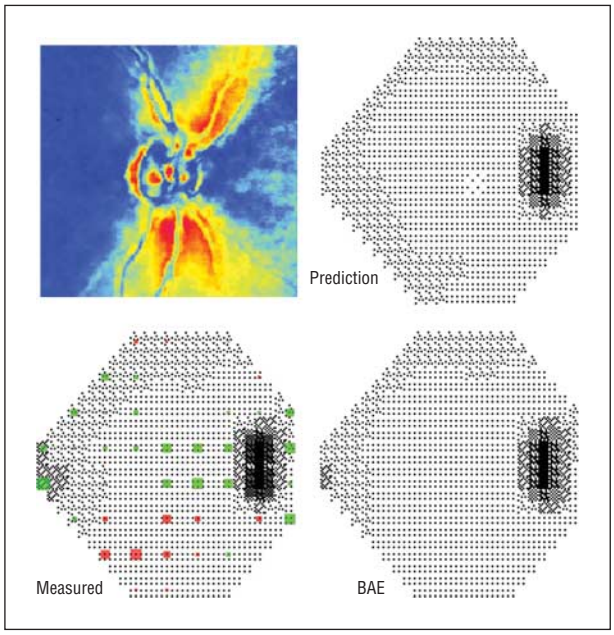


Figure 4. An example of good structure-function concordance (structure-function discordance index score of 0.15). The GDx VCC retinal nerve fiber layer thickness predicted the early nasal defect that was also indicated in measured and best-available estimate (BAE) visual fields. The structure-function pattern discordance map superimposed on the measured field shows small discordance flags (a green square indicates that the structural measure did not predict as low a visual sensitivity as the actual test result [discordance >0], while a red square means that the structural measure predicted a greater visual defect than was found in visual field testing [discordance <0]).

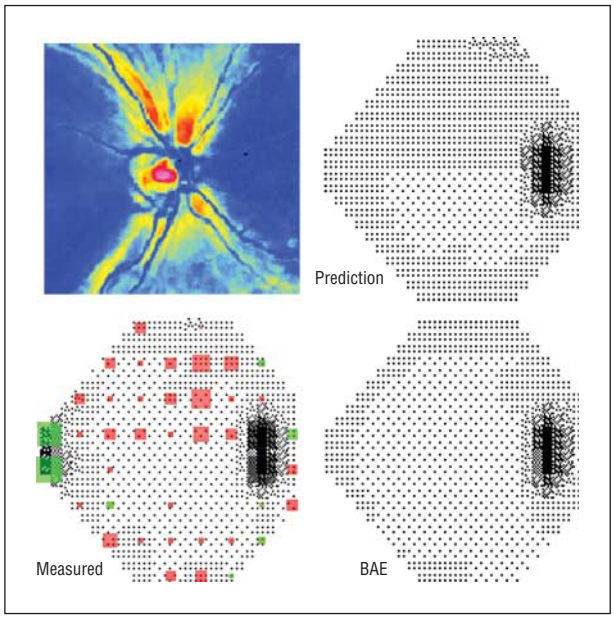


Figure 5. An example of satisfactory overall structure-function concordance (structure-function discordance index score of 0.28) but with local discordance. The retinal nerve fiber layer thickness did not predict the nasal defect in the measured visual field, which was flagged by the large green squares in the structure-function pattern discordance map (a green square indicates that the structural measure did not predict as low a visual sensitivity as the actual test result [discordance >0], while a red square means that the structural measure predicted a greater visual defect than was found in visual field testing [discordance <0]). BAE indicates best-available estimate.

The last example, with high discordance (SFDI score of 0.66), is demonstrated in Figure 7. The overall discordance was flagged across the entire VF in the structure-

Downloaded from www.archophthalmol.com at , on December 27, 2011
©2011 American Medical Association. All rights reserved.

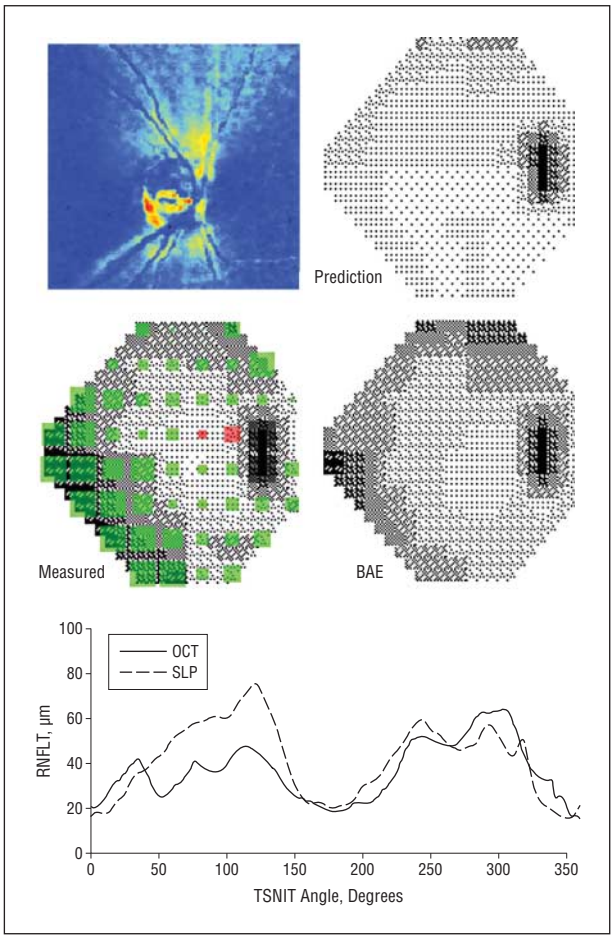


Figure 6. An example of substantial structure-function discordance (structure-function discordance index score of 0.58). The retinal nerve fiber layer thickness (RNFLT) prediction overestimated the measured visual field, especially in the inferior hemifield nasally, as flagged by the pattern discordance map (a green square indicates that the structural measure did not predict as low a visual sensitivity as the actual test result [discordance >0], while a red square means that the structural measure predicted a greater visual defect than was found in visual field testing [discordance <0]). The temporal-superior-nasal-inferior-temporal (TSNIT) RNFLT profile measured by optical coherence tomography (OCT) (appropriately rescaled[16]) and GDx VCC scanning laser polarimetry (SLP) were compared. BAE indicates best-available estimate.

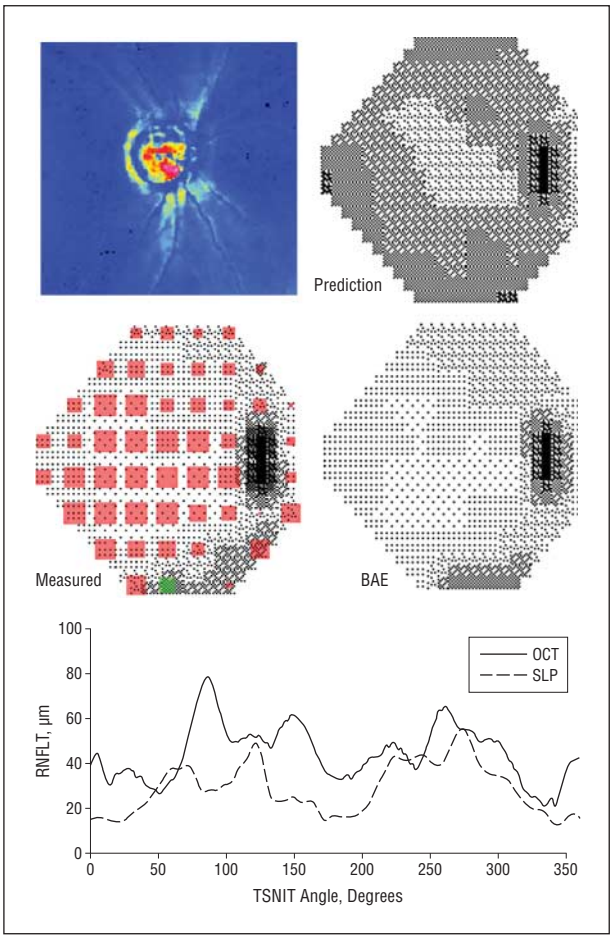


Figure 7. An example of significant structure-function discordance (structure-function discordance index score of 0.66). The retinal nerve fiber layer thickness (RNFLT) prediction significantly underestimated the entire measured visual field, as flagged by the pattern discordance map (a green square indicates that the structural measure did not predict as low a visual sensitivity as the actual test result [discordance >0], while a red square means that the structural measure predicted a greater visual defect than was found in visual field testing [discordance <0]). The temporal-superior-nasal-inferior-temporal (TSNIT) RNFLT profile measured by optical coherence tomography (OCT) (appropriately rescaled[16]) and GDx VCC scanning laser polarimetry (SLP) were compared. BAE indicates best-available estimate.

function PDM. The single measured VF was closer to the BAE VF, showing that the discordance was more likely to be caused by the GDx VCC RNFLT measurement.

COMMENT

Clinicians are expected to use both structural and functional assessment of the optic nerve to evaluate and monitor patients with glaucoma.[17] Quantitative tools for measuring visual function (for instance, SAP) and aspects of optic nerve anatomy such as the RNFL (for instance, the GDx VCC) are in common clinical use, but to our knowledge there have been no analysis algorithms available to describe the level of agreement between these different modes for the assessment of retinal ganglion cell integrity. The technique we describe addresses this clinical need.

The BRBF technique maps the structural measurement into the same domain as the functional measurement so that the clinicians can consider both measurements in tandem. The SFDI and structure-function PDM are proposed as tools to quantify and visualize any disagreement between the 2 types of measurements. Just as with a repeated VF test, the clinician should not expect exact agreement; however, serious disagreement may indicate that one or the other test was faulty, and sometimes the tests may need to be repeated.

The use of the predictability of the VF from the RNFLT by the BRBF for the discordance calculation played a key role in this method. **Figure 8** demonstrates that any regional variation of discordance within the VF can be accounted for by sensitivity values at various locations. Figure 8A illustrates the average distribution of VF sensitivity in the patients with glaucoma, and Figure 8B illustrates the distribution of mean absolute prediction errors; there is a clear regional variation that is similar to the distribution of sensitivity. In Figure 8C, the mean absolute (unsigned) structure-function discordance is cal-

Downloaded from www.archophthalmol.com at , on December 27, 2011
©2011 American Medical Association. All rights reserved.

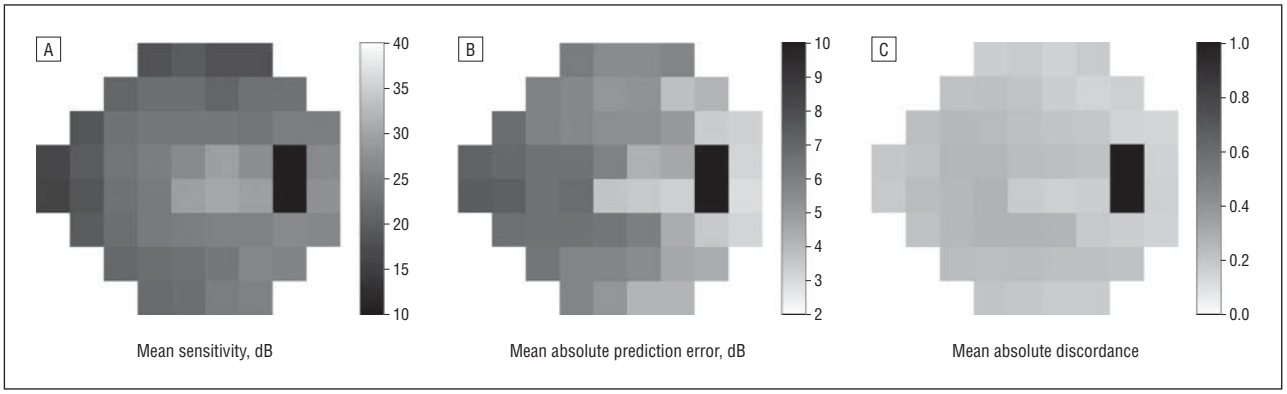


Figure 8. Illustration of spatial predictability. In this glaucomatous sample, the mean sensitivity (A), mean absolute prediction error (B), and mean absolute structure-function discordance (C) are shown for each visual field location with a gray scale.

culated at each location in the VF. Unsigned structure-function discordance normalizes the absolute prediction errors in decibels to be the unsigned percentile difference according to the VF predictability at different levels of measured VF sensitivity (Figure 2). The use of predictability based on prediction percentiles attenuates the regional variation of the structure-function discordance, demonstrating that the prediction errors have been normalized with respect to the measured VF sensitivity. Therefore, the discordance value was less affected by the change in predictability caused by the level of VF sensitivity and tended to reflect the intrinsic difference between the structural and functional measurements.

The most important potential use of the PDM and SFDI is in alerting the clinician to the possibility of suboptimal data such as those shown in Figure 5, Figure 6, and Figure 7. Suboptimal data may be in either the functional or structural measurement. In the functional domain, errors in quantifying neural loss may be systematic, such as the bias introduced by media opacity, or random, such as imprecision introduced by learning affects, fatigue, false-positive responses, and lens rim artifacts. In the structural domain, systematic error may be introduced by factors such as atypical scan pattern in GDx VCC images and measurement differences induced by variation in axial length; imprecision may be introduced by factors such as vitreous opacities, a poor tear film, poor fixation, and other known and unknown factors. Almost all sources of error will result in less than optimal concordance between structural and functional measurements, and the SFDI and PDM provide an easily interpretable flag to suboptimal data.

The PDM in Figure 5 shows local discordance in the nasal area and indicates relative underestimation of the sensitivity in the measured VF. The structure-predicted and measured VFs were compared with the reference BAE VF, which showed that this discordance was caused by test variability in the measured VF. The moderately sized red squares in the upper hemifield indicate that possible early RNFL loss has not manifested as VF loss in this patient.

Overall discordance is alerted by moderate SFDI scores and visual warning (large discordance squares) in the inferior part of the PDM in Figure 6. The OCT measurement of RNFLT was used to validate the RNFLT measurement by GDx VCC. Owing to measurement scaling differences between GDx VCC and OCT, the OCT measurements were adjusted by the scaling factors described by Leung et al.[16] The overestimation of RNFLT by GDx VCC when compared with the OCT measurement was apparent especially in the superior region and in this case may be caused by the atypical scan pattern in the GDx VCC RNFLT map.

Discordance across the entire VF is indicated by the SFDI and PDM in Figure 7. The agreement between the single and BAE VFs shows that the inaccurate RNFLT measurement made by GDx VCC was more likely to be the cause of the discordance. This was confirmed by the OCT measurement indicating significantly thicker RNFLT than that from GDx VCC.

In the absence of a better reference standard, BAE VF is used here as a means to compare how closely a measured VF resembles the presumed true field and similarly how the field predicted from structural data compares to the presumed true field. Although the BAE VF is derived from the measured VFs in these subjects and is therefore biased to resemble any of the measured VFs, it was found that in some cases the structural data predicted a worse VF and sometimes the measured VF was worse than the BAE VF. Comparison of the mean difference of the structure-predicted and measured VFs from the BAE VF showed that on average the measured VF was closer to the BAE VF. However, this may be, at least in part, bias from the fact that the BAE VF was determined from measured VFs themselves. Simply put, a VF from the 5 VFs composing the BAE VF is likely to be close to the BAE VF because of the method for calculating the BAE VF. The structure-predicted VF can be compared with the measured VF to score the structure-function discordance without the BAE VF reference used for this scientific study. It should be emphasized that the construction of a BAE VF test from available replicate testing was used in this study for illustrative purposes to demonstrate features of the method but is not necessary when the method is applied in a clinical setting or in other scientific studies.

The BRBF model was derived from measurements fulfilling strict reliability criteria (using VF reliability indices and GDx VCC image quality scores).[1] Similar reliability criteria were imposed on the measurements included in this study. Therefore, discordance quanti-

Downloaded from www.archophthalmol.com at , on December 27, 2011
©2011 American Medical Association. All rights reserved.

fied with measurements where these reliability criteria are not fulfilled would in turn be expected to be less reliable, although this hypothesis has not been tested in this work. However, the reliability criteria used are similar to those recommended by the instrument manufacturers and therefore represent the quality standards that should be aimed for.

The principles reported in this study could be developed for any structural measurement. The various commercial products of other imaging techniques (such as spectral-domain OCT) need to be evaluated independently. Furthermore, there is already a new version of the scanning laser polarimeter (GDx ECC; Carl Zeiss Meditec), which is likely to have a better prediction accuracy compared with the GDx VCC used in this study.

In conclusion, we describe for the first time a method that quantifies discordance between structural and functional measurements of glaucoma damage. The analysis is presented in a novel and clinically useful way such that high discordance may alert clinicians to an instance of poor-quality test data and low discordance may be reassuring that structural and functional findings match.

Submitted for Publication: November 5, 2010; final revision received February 11, 2011; accepted March 13, 2011.
Published Online: May 9, 2011. doi:10.1001/archophthalmol.2011.112
Author Affiliations: Department of Optometry and Visual Science, City University London (Drs Zhu, Crabb, and Garway-Heath) and National Institute for Health Research Biomedical Research Centre for Ophthalmology, Moorfields Eye Hospital NHS Foundation Trust and University College London Institute of Ophthalmology (Drs Zhu and Garway-Heath), London, England; Centre Universitaire d'Ophtalmologie, Centre d'excellence sur le vieillissement de Québec, Centre de Recherche, Fonds de la Recherche en Santé du Québec du Centre Hospitalier Affilié Universitaire de Québec, Department of Ophthalmology, Université Laval, Quebec City, Quebec, Canada (Dr Fredette); and Bascom Palmer Eye Institute, University of Miami Miller School of Medicine, Miami, Florida (Drs Fredette and Anderson).
Correspondence: David P. Crabb, PhD, Department of Optometry and Visual Science, City University London, Northampton Square, London EC1V 0HB, England (david.crabb.1@city.ac.uk).
Financial Disclosure: Dr Crabb received unrestricted financial support from Pfizer. Dr Anderson is a consultant for Carl Zeiss Meditec and Pfizer. Dr Garway-Heath is a consultant for and received financial support from Carl Zeiss Meditec and Pfizer. Dr Fredette received financial support from Pfizer.
Funding/Support: This work was chiefly supported by Pfizer with an investigator-initiated research grant and by funding from the Department of Health's National Institute for Health Research Biomedical Research Centre for Ophthalmology, Moorfields Eye Hospital NHS Foundation Trust and University College London Institute of Ophthalmology. Dr Garway-Heath's chair at University College London is supported by funding from the International Glaucoma Association. Data management was supported in part by a grant from the Fonds de la Recherche en Santé du Québec. Data acquisition was supported in part by a fellowship scholarship from Laval University, unrestricted donations from Carl Zeiss Meditec Humphrey, Welch Allyn, and Allergan, core grant P30 EY014801 from the National Eye Institute, and an unrestricted grant from Research to Prevent Blindness to the Bascom Palmer Eye Institute.
Role of the Sponsors: The sponsors and funding organizations had no role in the design or conduct of this research, data analysis, or writing of the manuscript.
Disclaimer: The views expressed in this article are those of the authors and not necessarily those of the Department of Health.

REFERENCES

1. 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(11):5657-5666.
2. Anderson RS. The psychophysics of glaucoma: improving the structure/function relationship. *Prog Retin Eye Res.* 2006;25(1):79-97.
3. Brigatti L, Caprioli J. Correlation of visual field with scanning confocal laser optic disc measurements in glaucoma. *Arch Ophthalmol.* 1995;113(9):1191-1194.
4. Weinreb RN, Shakiba S, Sample PA, et al. Association between quantitative nerve fiber layer measurement and visual field loss in glaucoma. *Am J Ophthalmol.* 1995;120(6):732-738.
5. Iester M, Mikelberg FS, Courtright P, Drance SM. Correlation between the visual field indices and Heidelberg retina tomograph parameters. *J Glaucoma.* 1997; 6(2):78-82.
6. Teesalu P, Vihanninjoki K, Airaksinen PJ, Tuulonen A, Läärä E. Correlation of blue-on-yellow visual fields with scanning confocal laser optic disc measurements. *Invest Ophthalmol Vis Sci.* 1997;38(12):2452-2459.
7. Garway-Heath DF, Holder GE, Fitzke FW, Hitchings RA. Relationship between electrophysiological, psychophysical, and anatomical measurements in glaucoma. *Invest Ophthalmol Vis Sci.* 2002;43(7):2213-2220.
8. Garway-Heath DF, Viswanathan A, Westcott M. Relationship between perimetric light sensitivity and optic disc neuroretinal rim area. In: Wall M, Wild JM, eds. *Perimetry Update 1998/1999.* The Hague, the Netherlands: Kugler Publications; 1999:381-389.
9. Anton A, Yamagishi N, Zangwill L, Sample PA, Weinreb RN. Mapping structural to functional damage in glaucoma with standard automated perimetry and confocal scanning laser ophthalmoscopy. *Am J Ophthalmol.* 1998;125(4):436-446.
10. Gardiner SK, Johnson CA, Cioffi GA. Evaluation of the structure-function relationship in glaucoma. *Invest Ophthalmol Vis Sci.* 2005;46(10):3712-3717.
11. Mai TA, Reus NJ, Lemij HG. Structure-function relationship is stronger with enhanced corneal compensation than with variable corneal compensation in scanning laser polarimetry. *Invest Ophthalmol Vis Sci.* 2007;48(4):1651-1658.
12. Bowd C, Zangwill LM, Medeiros FA, et al. Structure-function relationships using confocal scanning laser ophthalmoscopy, optical coherence tomography, and scanning laser polarimetry. *Invest Ophthalmol Vis Sci.* 2006;47(7): 2889-2895.
13. Schlottmann PG, De Cilla S, Greenfield DS, Caprioli J, Garway-Heath DF. Relationship between visual field sensitivity and retinal nerve fiber layer thickness as measured by scanning laser polarimetry. *Invest Ophthalmol Vis Sci.* 2004; 45(6):1823-1829.
14. 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(10):1809-1815.
15. Bishop CM. *Neural Network for Pattern Recognition.* Oxford, England: Oxford University Press; 1996.
16. Leung CK, Chan WM, Chong KK, et al. Comparative study of retinal nerve fiber layer measurement by StratusOCT and GDx VCC, I: correlation analysis in glaucoma. *Invest Ophthalmol Vis Sci.* 2005;46(9):3214-3220.
17. European Glaucoma Society. *Terminology and Guidelines for Glaucoma.* 3rd ed. Savona, Italy: European Glaucoma Society; 2008.

Downloaded from www.archophthalmol.com at , on December 27, 2011
©2011 American Medical Association. All rights reserved.