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

Haogang Zhu, MSc;[1,2] David P. Crabb, PhD;[1,*], Marie-Josée Fredette, MD;[3,4] Douglas R. Anderson, MD;[3] David F. Garway-Heath, MD [1,2]

1 Department of Optometry and Visual Science, City University London, London, UK

2 NIHR Biomedical Research Centre for Ophthalmology, Moorfields Eye Hospital NHS Foundation Trust and UCL Institute of Ophthalmology, London, UK

3 Bascom Palmer Eye Institute, University of Miami Miller School of Medicine, Miami, Florida, USA

4 Centre Universitaire d'Ophtalmologie (CUO), CEVQ, Centre de Recherche FRSQ du Centre Hospitalier Affilié Universitaire de Québec, Department of Ophthalmology, Université Laval, Quebec City, Canada

* Department of Optometry and Visual Science, City University London, Northampton Square, London, EC1V 0HB, UK

Email: david.crabb.1@city.ac.uk

Tel: 0044 (0)20 7040 0191

Disclosures

Professor David Crabb received unrestricted financial support from Pfizer, Inc., New York, NY, USA.

Professor Douglas Anderson is a consultant for Carl Zeiss Meditec, Inc., Dublin, CA, USA and for Pfizer, Inc., New York, NY, USA.

Professor David Garway-Heath is a consultant for and received financial support from Carl Zeiss Meditec, Inc., Dublin, CA, USA and for Pfizer, Inc., New York, NY, USA.

Professor Marie-Josée Fredette received financial support from Pfizer, Inc., New York, NY, USA.

Abstract

Objective: The visual field (VF) may be predicted from retinal nerve fibre layer thickness (RNFLT) using a Bayesian Radial Basis Function (BRBF). This study aimed to evaluate a new methodology to quantify and visualise discordance between structural and functional measurements in glaucomatous eyes.

Methods: Five GDxVCC RNFLT scans and 5 Humphrey SITA VF tests were obtained from 50 glaucomatous eyes from 50 patients. A best available estimate of the 'true' VF was calculated as the point-wise median of these 5 replications. This 'true' VF was compared with every single RNFLT-predicted VF from BRBF and every single measured VF. Predictability of VFs from RNFLT was established from previous data. A structure-function pattern discordance map (PDM) and structure-function discordance index (SFDI; values 0 to 1) were established from the predictability limits for each structure-function measurement pair to quantify and visualise the discordance between the structure-predicted and measured VFs.

Results: Mean absolute difference (MAD) between the structure-predicted and 'true' VFs was 3.9dB. MAD between single and 'true' VFs was 2.6dB. Mean of SFDI was 0.34 (SD 0.11). 39% of the structure-predicted VFs showed significant discordance (SFDI>0.3) from measured VFs.

Conclusions: BRBF, on average, predicts the 'true' VF from RNFLT slightly less well than a measured VF from the 5 VFs compromising the 'true' VF. The PDM highlights locations with structure-function discordance, with the SFDI 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.

Introduction

Glaucoma is a progressive optic neuropathy in which structural damage is evident at the optic nerve head (ONH) and the retinal nerve fibre 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, which, 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 structural loss [1]. However, the relationship between the structural and functional measurements is still not well understood, mainly due 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 ONH 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). SAP 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 art tools for the diagnosis and management of glaucoma.

When multiple types of measurement assessing glaucoma damage are available for the same patient, an important question is: "Are they consistent with each other, and if not, to what degree and why are they discordant?" Thinking about this question 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 ONH (Heidelberg Eye Explorer (HEYEX) version 1.7.0, Heidelberg Engineering, Heidelberg, Germany). This enables visualization of structure-

function agreement, or spatial concordance, at an ONH 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 (e.g., mean RNFLT in 6 sectors) and groups of, or individual points in the VF, and assess the curve linear (e.g., log-linear) or monotonic association between the two 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). Firstly, 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 GDxVCC [Carl Zeiss Meditec, Dublin, CA, USA]), or a grouping of visual field locations, so the spatial resolution of the measurements is lost. Secondly, 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. Thirdly, 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 anatomic variation, the structure-function relationship in these studies is reported as a statistical association based on measurements from a population as average typical relationship, 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 dataset used to establish the structure-function relationship.

In contrast to association-based methods, a more flexible non-linear prediction model, the Bayesian Radial Basis Function (BRBF), has been developed and validated to predict the VF from the RNFLT [1, 14]. The model was built and tested on the measurements of 535 eyes from 535 subjects from three independent centres. It was shown that the BRBF model, on average, can predict a patient's visual field from the RNFLT almost as well as another visual field from the patient. Moreover, the BRBF derived a high resolution structure-function relationship that is consistent with known typical location of retinal ganglion cells that project their axons toward particular meridians of the optic disc [15]. 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 utilise the BRBF method to predict the expected result of a VF test based on structural measurements. Thereafter, an analysis was undertaken to quantify and visualise the discordance between an individual VF test and the expectation based on structural information. This method is demonstrated, on a test-retest dataset, in hope of providing a basis for a clinical tool that provides an 'alert' for inconsistency of a VF test with structural measurements.

Methods

Subjects

A test-retest dataset 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 (SITA)-Standard program of the Humphrey Field Analyzer (HFA, Carl Zeiss Meditec Inc, Dublin, CA, USA). The dataset covers a wide spectrum of glaucoma severity from early to advanced stage, judged by their treating physicians to be under control with stability of nerve damage and VF loss. Inclusion criteria for these patients included: (1) 18 years old or more, (2) controlled intraocular pressures, (3) best corrected visual acuity of 20/40 or better, (4) less than 5 diopters of spherical, 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 non-glaucomatous structural or functional abnormality, (7) no history of amblyopia, (8) mental and physical incapacity to perform the tests, and (9) willingness to participate in the study. Only one 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 (even distribution of selected eyes across the mean deviation).

Each subject attended five sessions within an eight-week period. During each of the five sessions, the subjects had SAP, scanning laser polarimetry (SLP) and optical coherence tomography (OCT) measurements done. SAP was performed with the 24-2 SITA-Standard program of the HFA with the Goldmann size III stimulus. The SAP test used to establish the diagnosis and eligibility was not included in the 5 repeat tests obtained for the study experiments. The SLP measurements were taken with the GDxVCC. All scans were acquired through undilated pupils. In this dataset, measurement reliability indices were similar to that of the dataset used to establish the BRBF model [1]: for VF tests, fixation losses $\leqslant$20%, false-positive rate $\leqslant$20% and false-negative rate $\leqslant$33%, and for for GDxVCC images, image

quality score ≥7 and typical scan score ≥80. OCT images were taken with the Stratus OCT (Carl Zeiss Meditec Inc., Dublin, CA, USA) using the 3.4mm fast scan protocol.

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 anonymised and transferred to a secure database held at City University London for the conduct of the present analysis.

Analysis

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

Structure-predicted VF

The structure-function relationship was previously modelled with a BRBF which was used to predict the VF from the RNFLT profile from the GDxVCC. The model had been trained on RNFLT and VF measurements of 229 eyes, and derived a structure-function relationship that is consistent with the retinal anatomy such that the structure-predicted VF can be understood as the RNFLT representation in the VF domain [1, 14]. The structure-function discordance, as described below, is defined as the 'difference' between this structure-predicted and the measured VF.

Best available estimate of 'true' VF

In order to compare the predicted and measured VFs, the true VF should be known, but high variability of testing means it cannot be known from any particular VF test. Therefore, a best available estimate of the true VF was calculated as the point-wise median of the 5 repeated VFs. It is then assumed that the best available estimate VF provides closer estimate of the true VF than any single measurement. The best available estimate VF will be referred as the 'true' VF in the subsequent description. The 'true' 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 VFs were discordant. However, the methodology proposed in this study is not constrained by the availability of a 'true' VF.

The difference between each structure-predicted VF and the 'true' VF was quantified. This difference was compared with the difference between each single VF from the 5 test-retest set and the 'true' VF.

VF predictability and quantification of structure-function discordance

The discordance between the structure-predicted VF and measured VF may be easily calculated as the point-wise difference of the two sensitivities. However, the predictability of the VF varies with the VF sensitivity itself. In Figure 2(a), as reported previously [1, 14], 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 predictability can be quantified as the percentile of the predicted VF sensitivity for any given measured VF sensitivity. This maps the VF sensitivity in dB to a prediction percentile ranging from 0 to 1. The percentile curves for the predicted VF sensitivity are illustrated in Figure 2(b) at the measured VF sensitivities

of 5dB, 15dB and 25dB. The variable predictability at different measured sensitivities is illustrated by the different shapes of the percentile curves in Figure 2(b): the corresponding curve shifts towards the right and has a steeper slope (increases from 0 to 1 more rapidly) at higher measured sensitivity. Because the predictability profile is built from the BRBF model that was previously constructed from measurements selected with strict quality criteria [1], the influence of unreliable measurement on the estimate of predictability is assumed to be minimal.

Due to the greater predictability at high sensitivity, any given predicted/measured difference 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 sensitivity 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 5dB predicted/measured difference at a measured sensitivity of 5dB, the discordance between a 0dB prediction and 5 dB measurement is 5.1% (on the 5dB predictability curve; Figure 2(b)), but for a 5dB difference at a 15dB measured sensitivity, the discordance between 10dB prediction and 15dB measurement is significantly higher at 17.3% (on the 15dB predictability curve; Figure 2(b)). Similarly, the discordance between 20dB prediction and 25dB measurement is even higher: 30.1%. Therefore, by using the predictability percentiles, the difference between the structure-predicted and measured VF sensitivity in dB is normalised with respect to the predictability at any given level of measured sensitivity.

The signed point-wise discordance between the structure-predicted and measured VF sensitivity 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 (PDM) and structure-function discordance index (SFDI)

The point-wise structure-function discordance can be visualised with a novel tool, named the structure-function pattern discordance map (PDM), with a technique similar to the Hinton diagram, which has been used to visualise the weight parameters in a multilayer percepteon neural network [16]. As shown in Figure 1, the structure-function PDM is a 24-2 greyscale 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 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 HFA VF chart to describe the deviation from the normative database, the PDM provides an easily-understood clinical tool to identify the point-wise deviation of the structure-predicted VF from the measured VF.

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

The SFDI and the PDM were used on the test-retest dataset to evaluate the structure-function discordance in this glaucomatous sample. The ‘true’ 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.

Result

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

The SFDI was calculated for each of the 5 pairs of structural and functional measurements taken in the same session. In this glaucomatous sample, the mean SFDI was 0.34 (SD 0.11), and the distribution of SFDI is given by the histogram in Figure 3. Although SFDI is a continuous variable, with a suggestion of bimodal distribution, an arbitrary threshold (SFDI=0.3) was chosen to divide the measurements into two groups: a group with good overall concordance and the other with noticeable discordance. Sixty-one percent of the measurements fall in the group with good concordance. For the group of measurements with good overall structure-function concordance (SFDI≤0.3), the mean SFDI was 0.25 (SD of 0.04) compared with the SFDI (0.41±0.08) for the group with noticeable discordance (SFDI>0.3).

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

In Figure 4, the structure-predicted, measured and 'true' VFs all indicated an early defect in the nasal, nasal-superior and nasal-inferior areas. The SFDI 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 it may fail to pick up local discordance. For example, the SFDI 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 two large green 'discordance squares' in the nasal area.

Figure 6 is an example with moderate overall discordance (SFDI=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 'true' VF, showing that the structural measurement is more likely to be the cause of the discordance.

The last example, with high discordance (SFDI=0.66) is demonstrated in Figure 7. The overall discordance was flagged across the whole VF in the structure-function PDM. The single measured VF was closer to the 'true' VF and showed that the discordance was more likely to be caused by the GDxVCC RNFLT measurement.

Discussion

Clinicians are expected to use both structural and functional assessment of the optic nerve to evaluate and monitor glaucoma patients [17]. Quantitative tools for measuring visual function (for instance, SAP) and aspects of optic nerve anatomy, such as the RNFL (for instance, the GDxVCC), are in common clinical usage, but, as yet, 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 described in this paper 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 visualise any disagreement between the two different types of measurements. Just as with a repeat visual field test, the clinician should not expect exact agreement, but 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 8(a) illustrates the average distribution of VF sensitivity in the glaucoma patients and Figure 8(b) illustrates the distribution of mean absolute prediction errors; there is a clear regional variation which is similar to the distribution of sensitivity. In Figure 8(c), the mean unsigned structure-function discordance is calculated at each location in the VF. Unsigned structure-function discordance normalises the absolute prediction errors in dB 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 normalised 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 is in alerting the clinician to the possibility of 'bad data' such as those shown in Figure 5, 5 and 6. Bad data may be either in 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, lens rim artifacts, and many other factors. In the structural domain, systematic error may be introduced by factors such as atypical scan pattern in GDxVCC 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 'true' VF, which showed that this discordance was caused by test variability in the measured VF. The moderate 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 and visual warning (large 'discordance squares') in the inferior part of PDM in Figure 6. The OCT measurement of RNFLT was used to validate the RNFLT measurement by GDxVCC. Due to measurement scaling differences between GDxVCC and OCT, the OCT measurements were adjusted by the scaling factors described by Leung et al [18]. The overestimation of RNFLT by GDxVCC, 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 GDxVCC RNFLT map.

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

In the absence of a better reference standard, 'true' 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 'true' VF is derived from the measured VFs on 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 test was worse than the 'true' field. Comparison of the mean difference of the structure-predicted and of the measured VFs from the 'true' VF showed that, on average, the measured VF was closer to the 'true' VF. However, this may be, at least in part, bias from the fact that the 'true' VF was determined from measured VFs themselves. Simply put, a VF from the 5 VFs comprising the 'true' VF is likely to be close to the 'true' VF because of the method for calculating the 'true' VF. The structure-predicted VF can be compared with the measured VF to score the structure-function discordance without the 'true' VF reference used for this scientific study. So it should be emphasised that the construction of a 'true' 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 predictability of the VF from the RNFLT was derived from previously developed BRBF model [1, 14]. As it was discussed in Zhu et al [1], despite of the general similarity between the prediction limits and VF test-retest limits, predictions at the normal end of the range still tend to be lower than the VF measurements, and at the damaged end, predictions tend to be higher than VF measurements (Figure 2). This likely reflects the difficulty the prediction from RNFLT images have in identifying small, focal defects. Moreover, the 'floor effect' in the VFs and GDxVCC SLP images [19, 20] and the atypical scan pattern in GDxVCC SLP images, which may be associated with glaucoma severity [21], may be additional causes of the overestimation at the lower end of the VF sensitivity. The tendency of overestimated and

underestimated prediction may bias the estimate of predictability, causing more overestimation discordance at the normal end and more underestimation discordance at the damaged end. Therefore, improving the BRBF model may gain better estimate of predictability profile that will lead to a more robust method. Potential improvement can be made by using alternative imaging techniques for structural measurement, by involving larger datasets with more variety of local defect, and technically, by improvement of BRBF inference algorithm such as the modification using a more realistic non-Gaussian noise in the prediction model.

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 SLP (GDxECC, Carl Zeiss Meditec, Dublin, CA, USA), which is likely to have a better prediction accuracy compared to the GDxVCC used in this study.

In conclusion, this paper presents, for the first time, a methodology 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.

Acknowledgements

This work was chiefly supported by Pfizer with an Investigator Initiated research grant, and funding from the Department of Health's NIHR Biomedical Research Centre for Ophthalmology at Moorfields Eye Hospital and UCL Institute of Ophthalmology. The views expressed in the publication are those of the authors and not necessarily those of the Department of Health. Professor Garway-Heath´s chair at UCL is supported by funding from the International Glaucoma Association. Data management was supported in part by a grant from the FRSQ (Fond de Recherche en Santé du Québec). Data acquisition was supported in part by a fellowship scholarship from Laval University, Quebec City, Quebec; and unrestricted donations from Carl Zeiss Meditec Humphrey, Dublin, California; Welch Allyn, Skaneateles, New York; Allergan, Inc., Irvine, California; a Core Grant P30 EY014801 awarded by the National Eye Institute, Bethesda, Maryland; and an unrestricted grant from Research to Prevent Blindness, Inc., New York, New York to the Bascom Palmer Eye Institute. The sponsors and funding organizations had no role in the design or conduct of this research, data analysis, or writing of the manuscript.

References

1. Anderson RS. The psychophysics of glaucoma: Improving the structure/function relationship. *Prog Ret Eye Res* 2006;25:79-97

2. Brigatti L, Caprioli J. Correlation of visual field with scanning confocal laser optic disc measurements in glaucoma. *Arch Ophthalmol* 1995;113(9):1191-1194.

3. 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;120:732-738.

4. 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.

5. Teesalu P, Vihanninjoki K, Airaksinen P, Tuulonen A, Laara E. Correlation of blue-on-yellow visual fields with scanning confocal laser optic disc measurements. *Invest Ophthalmol Vis Sci* 1997;38:2452-2459.

6. Garway-Heath DF, Holder GE, Fitzke FW, Hitchings RA. Relationship between Electrophysiological, Psychophysical, and Anatomical Measurements in Glaucoma. *Invest Ophthalmol Vis Sci* 2002;43:2213-2220.

7. 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 Netherlands: Kugler Publications The Hague; 1999:381-389.

8. Anton A, Yamagishi N, Zangwill L, Sample P, Weinreb R. Mapping structural to functional damage in glaucoma with standard automated perimetry and confocal scanning laser ophthalmoscopy. *Am J Ophthalmol* 1998;125(4):436-446.

9. Gardiner SK, Johnson CA, Cioffi GA. Evaluation of the Structure-Function Relationship in Glaucoma. *Invest Ophthalmol Vis Sci* 2005;46:3712-3717.

10. 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;48:1651-1658.

11. Bowd C, Zangwill LM, Medeiros FA, et al. Structure and Function in Glaucoma: The Relationship between a Functional Visual Field Map and an Anatomic Retinal Map. *Invest Ophthalmol Vis Sci* 2006;47:2889-2895.

12. Schlottmann PG, Cilla SD, 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;45:1823-1829.

13. Zhu H, Crabb DP, Schlottmann PG, et al. Predicting Visual Function from the Measurements of Retinal Nerve Fibre Layer Structure. *Invest Ophthalmol Vis Sci* 2010;Accepted. In press.

14. Zhu H, Crabb DP, Garway-Heath DF. A Bayesian Radial Basis Function Model to Link Retinal Structure and Visual Function in Glaucoma. *The 3rd International Conference of Bioinformatics and Biomedical Engineering.* Beijing, China: IEEE; 2009:1-4.

15. 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.

16. Bishop CM. *Neural network for pattern recognition*: Oxford University Press; 1996.

17. Terminology and guidelines for glaucoma (3rd edition). Savona, Italy: European Glaucoma Society; 2008.

18. 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;46:3214-3220.

Figure legends

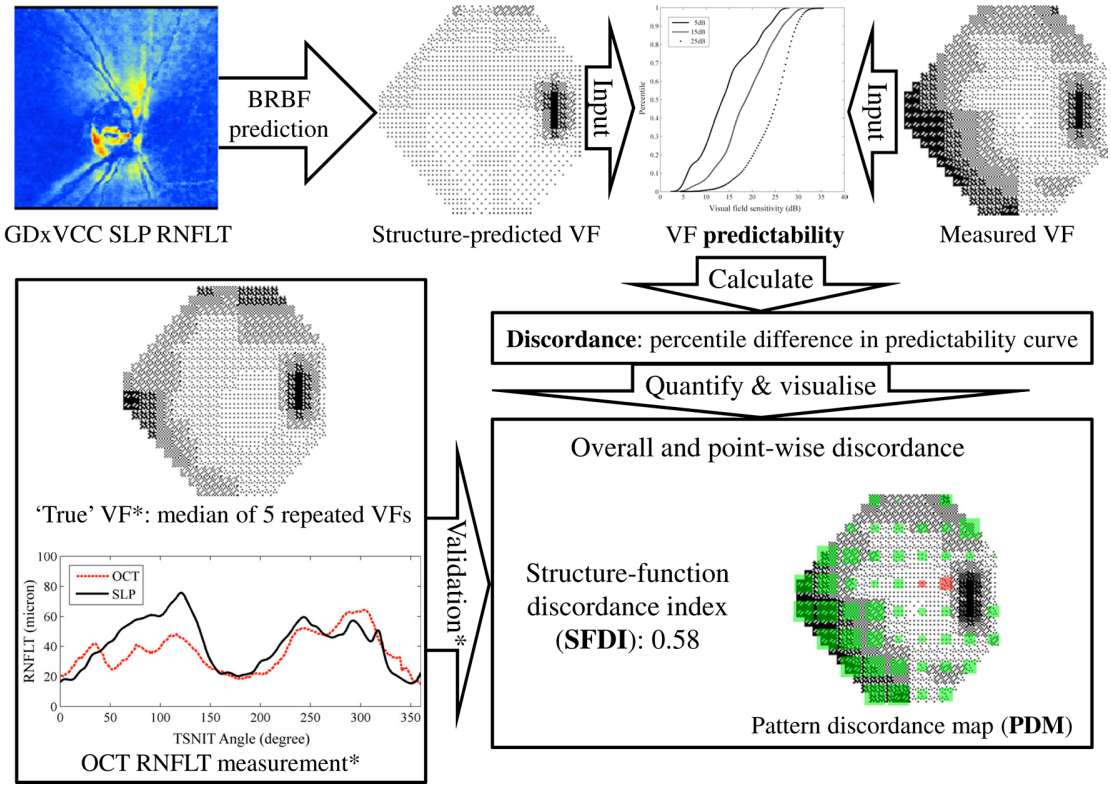

Figure 1. A flowchart summarising the analysis in the study. Note that the steps with star (*) symbols are only for the purpose of validating the method and are not necessary when using the proposed method in practice.

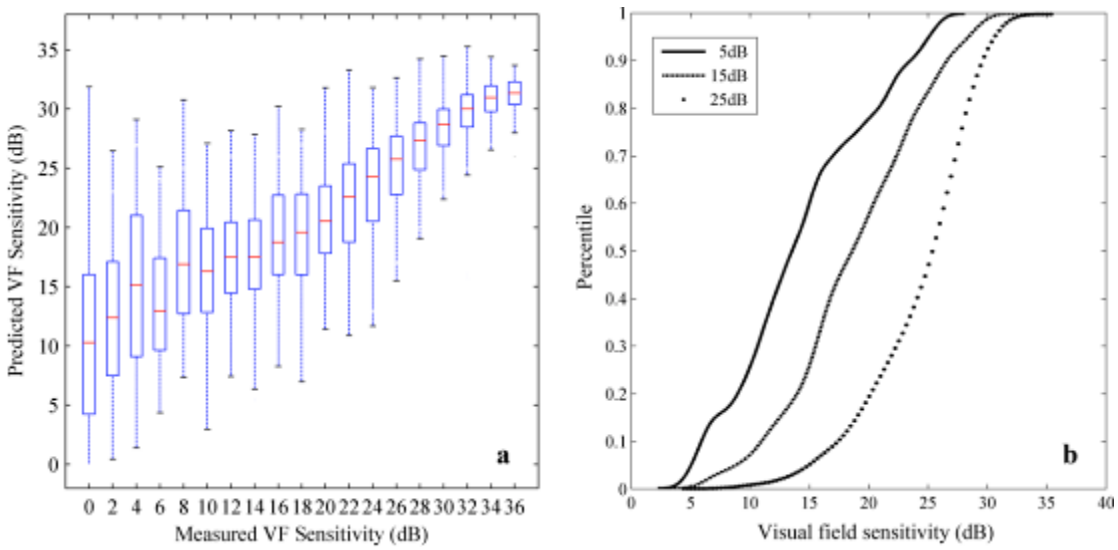

Figure 2. BRBF predictability and percentile curves. (a) BRBF prediction ranges (predictability) at different visual field sensitivities. Thin line tips and 'box' ends indicate 90% (5%—95%) and interquartile (25%—75%) prediction limits. (b) Percentile curves of predictability at measured sensitivities of 5dB, 15dB and 25dB.

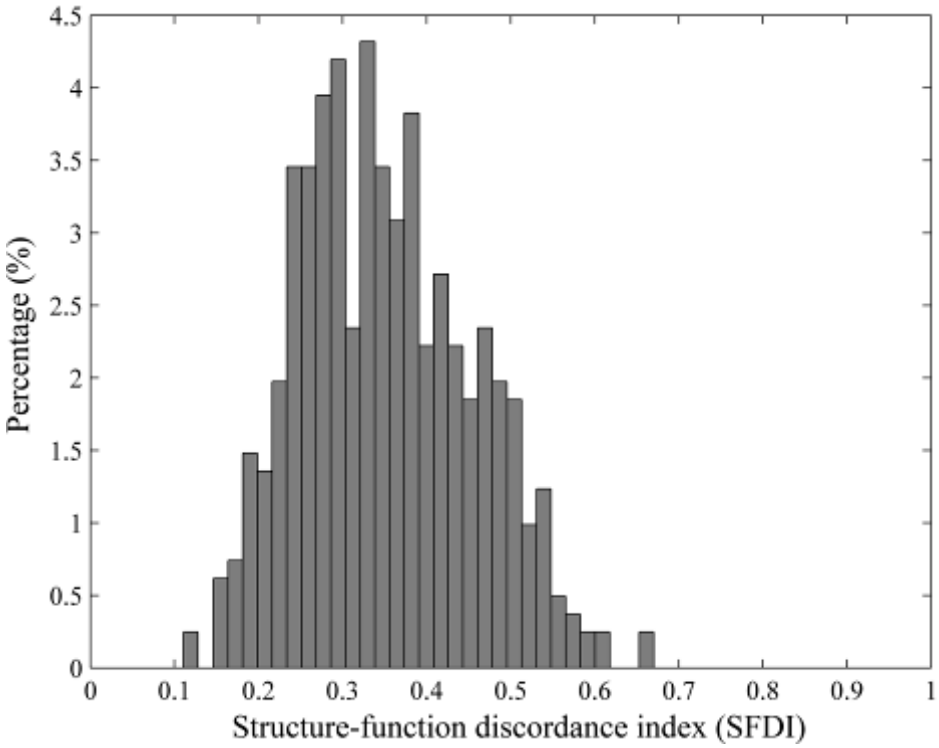

Figure 3. Histogram of structure-function discordance index (SFDI) in this test-retest sample of glaucomatous eyes.

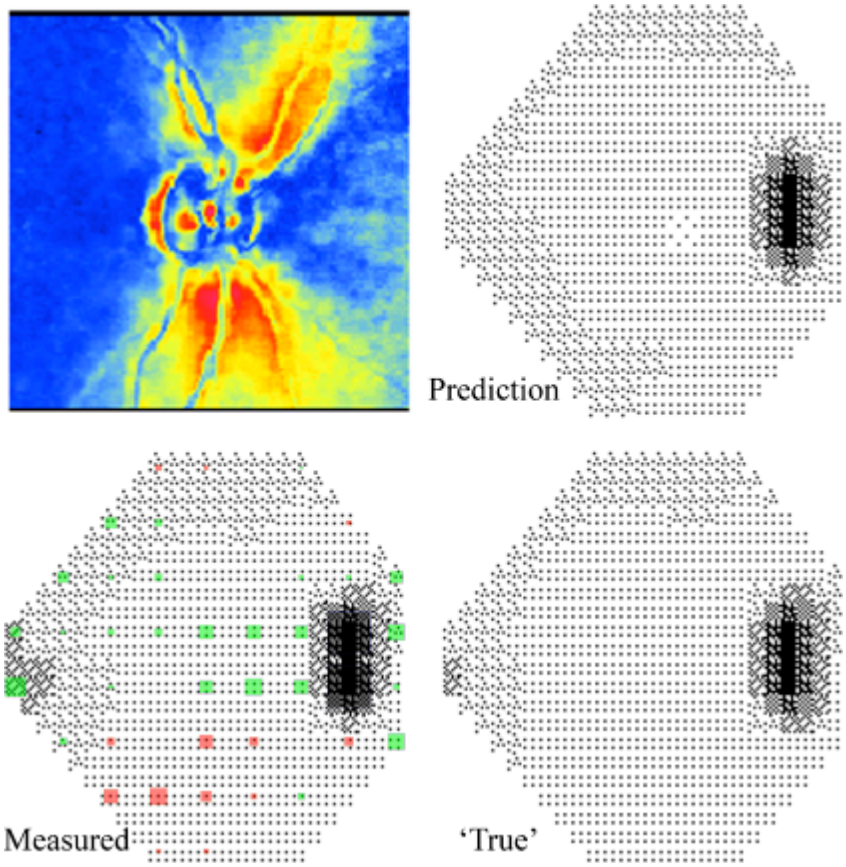

Figure 4. An example of good structure-function concordance (SFDI=0.15). The GDxVCC RNFL thickness predicted early nasal defect that was also indicated in measured and 'true' visual fields (VFs). Structure-function pattern discordance map (PDM) superimposed on the measured VF shows small discordance flags (green/red squares).

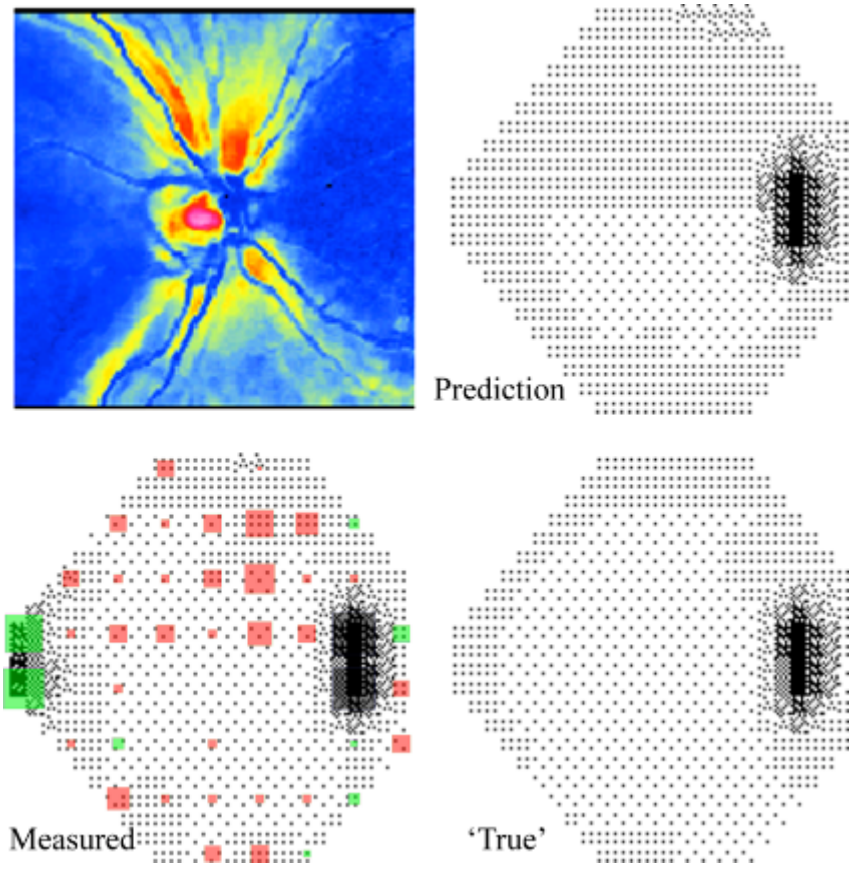

Figure 5. An example with satisfactory overall structure-function concordance (SFDI=0.28) but with local discordance. The RNFL thickness did not predict the nasal defect in the measured visual field, which was flagged by the large green squares in structure-function pattern discordance map.

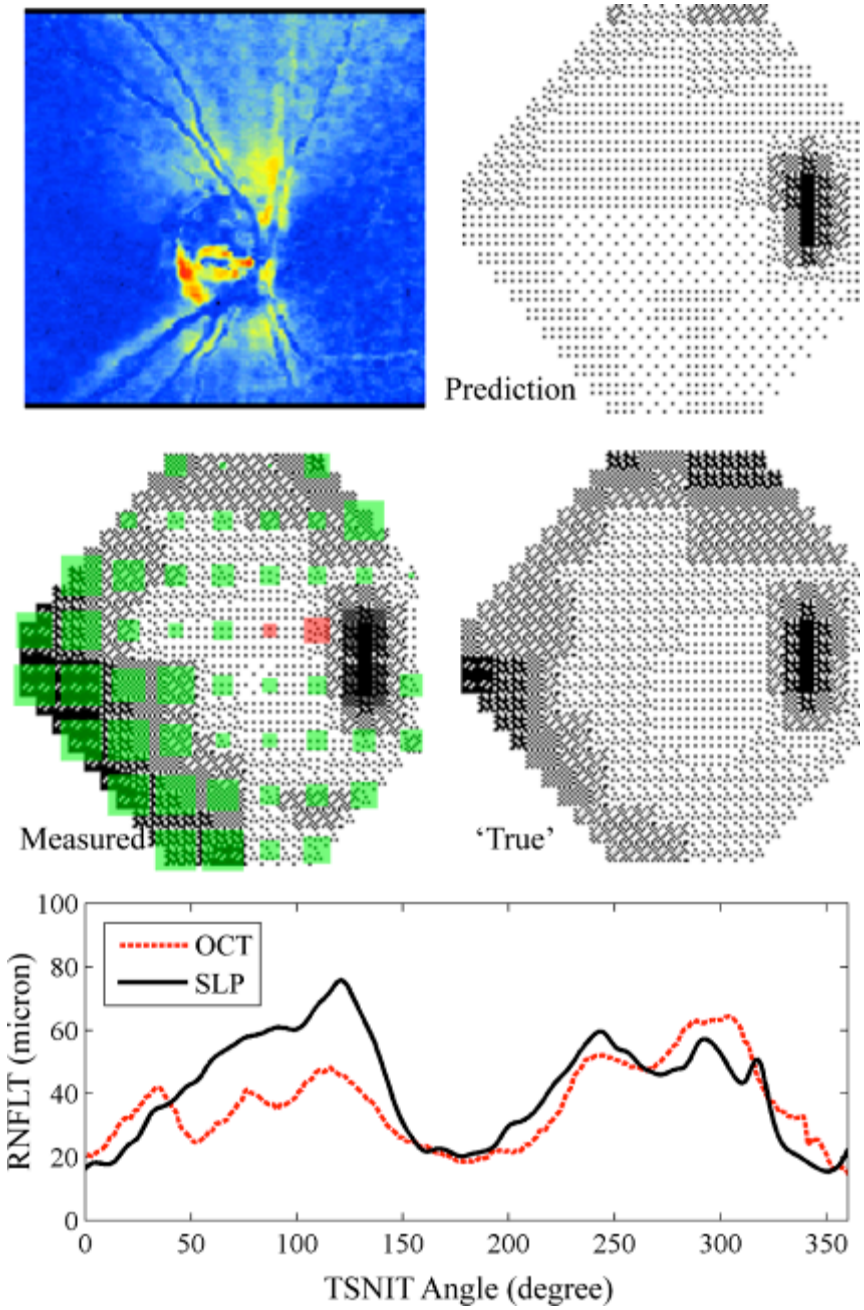

Figure 6. An example with substantial discordance (SFDI=0.58). The RNFL thickness prediction overestimated the measured visual field, especially in the inferior hemifield nasally, as flagged by the pattern discordance map. TSNIT RNFL thickness profile measured by OCT (appropriately rescaled [18]) and GDxVCC were compared.

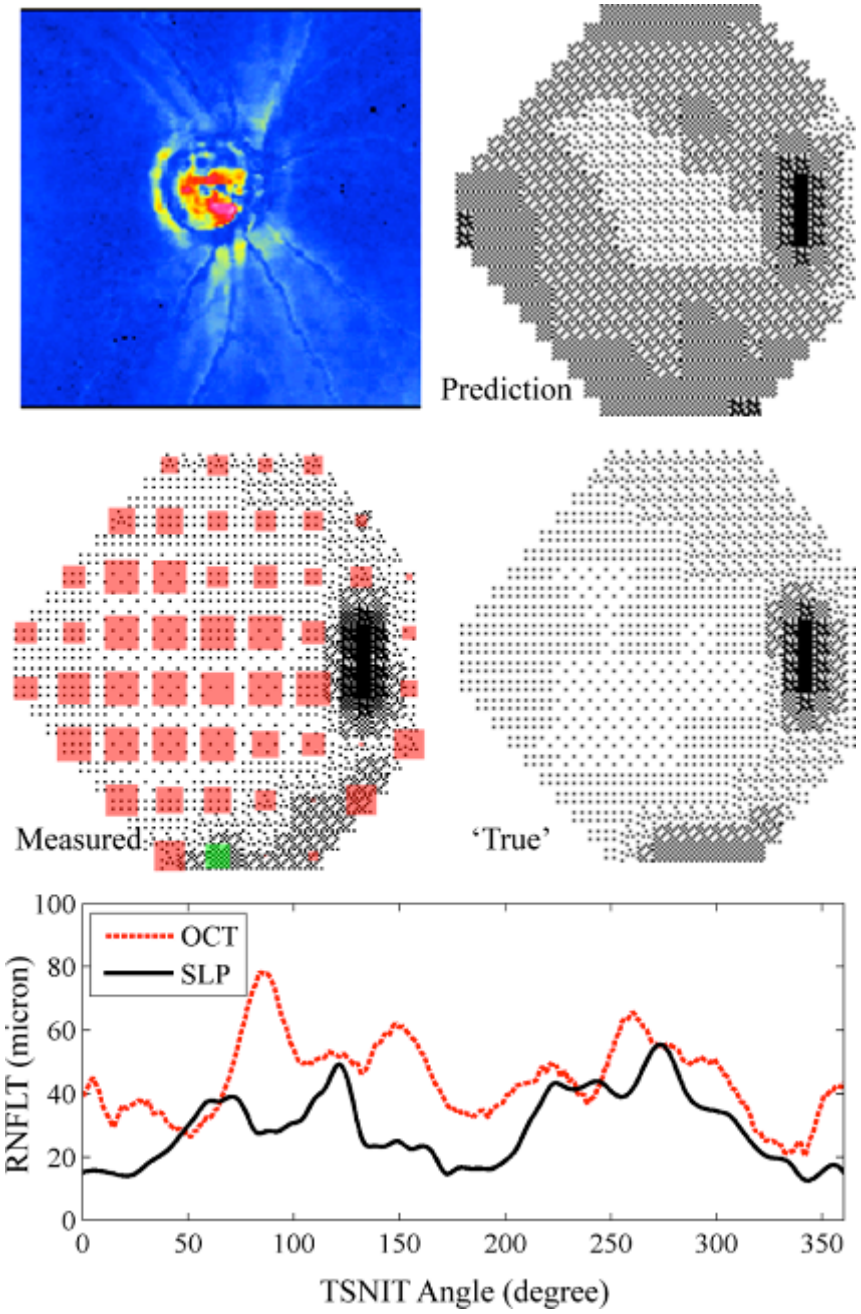

Figure 7. An example of significant structure-function discordance (SFDI=0.66). The RNFL thickness prediction significantly underestimated the whole measured visual field, as flagged by the pattern discordance map. TSNIT RNFL thickness profile measured by OCT (appropriately rescaled [18]) and GDxVCC were compared.

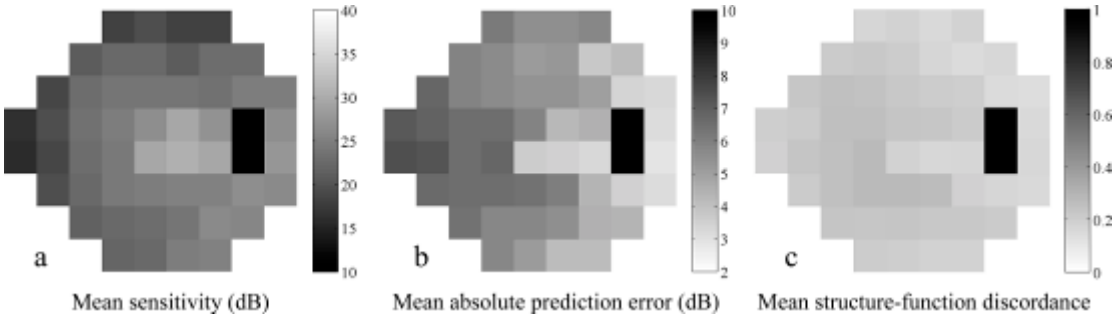

Figure 8. Illustration of spatial predictability. In this glaucomatous sample, the mean sensitivity (a), mean absolute prediction error (b) and mean absolute structure-function discordance (c) are shown for each visual field location with a grey scale.

1. Zhu H, Crabb DP, Schlottmann PG, et al. Predicting Visual Function from the Measurements of Retinal Nerve Fibre Layer Structure. *Invest Ophthalmol Vis Sci* 2010;51:5657-5666.

2. Anderson RS. The psychophysics of glaucoma: Improving the structure/function relationship. *Prog Ret Eye Res* 2006;25:79-97

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;120:732-738.

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 P, Tuulonen A, Laara E. Correlation of blue-on-yellow visual fields with scanning confocal laser optic disc measurements. *Invest Ophthalmol Vis Sci* 1997;38:2452-2459.

7. Garway-Heath DF, Holder GE, Fitzke FW, Hitchings RA. Relationship between Electrophysiological, Psychophysical, and Anatomical Measurements in Glaucoma. *Invest Ophthalmol Vis Sci* 2002;43:2213-2220.

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 Netherlands: Kugler Publications The Hague; 1999:381-389.

9. Anton A, Yamagishi N, Zangwill L, Sample P, Weinreb R. 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;46:3712-3717.

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;48:1651-1658.

12. Bowd C, Zangwill LM, Medeiros FA, et al. Structure and Function in Glaucoma: The Relationship between a Functional Visual Field Map and an Anatomic Retinal Map. *Invest Ophthalmol Vis Sci* 2006;47:2889-2895.

13. Schlottmann PG, Cilla SD, 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;45:1823-1829.

14. Zhu H, Crabb DP, Garway-Heath DF. A Bayesian Radial Basis Function Model to Link Retinal Structure and Visual Function in Glaucoma. *The 3rd International Conference of Bioinformatics and Biomedical Engineering.* Beijing, China: IEEE; 2009:1-4.

15. 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.

16. Bishop CM. *Neural network for pattern recognition*: Oxford University Press; 1996.

17. Terminology and guidelines for glaucoma (3rd edition). Savona, Italy: European Glaucoma Society; European Glaucoma Society 2008.

18. 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;46:3214-3220.

19. Wyatt HJ, Dul MW, Swanson WH. Variability of visual field measurements is correlated with the gradient of visual sensitivity *Vision Research* 2007;47:925-936.

20. Blumenthal EZ, Horani A, Sasikumar R, Garudadri C, Udaykumar A, Thomas R. Correlating Structure With Function in End-Stage Glaucoma. *Ophthalmic Surg Lasers Imaging* 2006;37:218-223.

21. Yanagisawa M, Tomidokoro A, Saito H, et al. Atypical retardation pattern in measurements of scanning laser polarimetry and its relating factors. *Eye* 2009;23:1796-1801.