



Contents lists available at ScienceDirect

Experimental Eye Research

journal homepage: www.elsevier.com/locate/yexerA geometric morphometric assessment of the optic cup in glaucoma[☆]Paul G. Sanfilippo^a, Andrea Cardini^b, Ian A. Sigal^c, Jonathan B. Ruddle^{a,d}, Brian E. Chua^{a,d}, Alex W. Hewitt^{a,d}, David A. Mackey^{a,e,f,*}^a Centre for Eye Research Australia, University of Melbourne, Australia^b Dipartimento di Biologia Animale, Università di Modena e Reggio Emilia, via Campi 213, Modena 41100, Italy & Hull York Medical School, The University of Hull, Cottingham Road, Hull HU6 7RX, UK^c Ocular Biomechanics Laboratory, Devers Eye Institute, Portland, Oregon, USA^d Royal Victorian Eye & Ear Hospital, Melbourne, Australia^e Discipline of Medicine, University of Tasmania, Hobart, Tasmania^f Lions Eye Institute, University of Western Australia, Centre for Ophthalmology and Visual Science, Perth, Australia

ARTICLE INFO

Article history:

Received 14 January 2010

Accepted in revised form 15 June 2010

Available online xxx

Keywords:

optic disc morphology

optic cup shape

shape analysis

geometric morphometrics

ovality

form-factor

discriminant analysis

ABSTRACT

The morphologic appearance of the optic disc is of interest in glaucoma. In contrast to descriptive classification systems that are currently used, a quantitative approach to the analysis of optic disc morphology is required.

Our goal was to determine the optimal method for quantifying optic cup shape by comparing traditional (ovality, form-factor and neuroretinal rim (NRR) width ratio) and geometric morphometric approaches. Left optic disc stereophotographs of 160 (80 normal and 80 glaucomatous (stratified by severity)) subjects were examined. The optic cup margins were stereoscopically delineated with a custom tracing system and saved as a series of discrete points. The geometric morphometric methods of elliptic Fourier analysis (EFA) and sliding semi-landmark analysis (SSLA) were used to eliminate variation unrelated to shape (e.g. size) and yield a series of shape variables. Differences in optic cup shape between normal and glaucoma groups were investigated. Discriminant functions were computed and the sensitivity and specificity of each technique determined. Receiver operator characteristic (ROC) curves were calculated for all methods and evaluated in their potential to discriminate between normal and glaucomatous eyes based on the shape variables. All geometric morphometric methods revealed differences between normal and glaucomatous eyes in optic cup shape, in addition to the traditional parameters of ovality, form-factor and NRR width ratio ($p < 0.0005$). SSLA (minimum bending energy criterion – 18 points) had the best sensitivity (83%) and area under the curve (AUC) (0.91). EFA (72 points) performed similarly well (74%, 0.89) as did the set of traditional shape-based variables (76%, 0.86). This study demonstrated that a geometric morphometric approach for discriminating between normal and glaucomatous eyes in optic cup shape is superior to that provided by traditional single parameter shape measures. Such analytical techniques could be incorporated into future automated optic disc screening modalities.

© 2010 Elsevier Ltd. All rights reserved.

1. Introduction

The morphologic appearance of the optic disc is paramount in evaluating its structural integrity. Various classification systems of

disc morphology have been proposed and utilize descriptive terms such as ‘focal ischaemic’, ‘senile sclerotic’, ‘shallow’ and ‘excavated’ to describe disc appearance. However, such clinical characterization is fundamentally subjective and a method of quantitative assessment of variation in disc morphology remains elusive. This is, in part, because measurement of shape is a more abstract notion than it may intuitively appear to be. Traditionally, morphologic analysis of the optic disc has been performed by calculating ratios of linear dimensions, such as ovality, or a parameter called form-factor (Gloster, 1975; Jonas et al., 1988a,b,d, 1997; Jonas and Papastathopoulos, 1996). The problem of shape analysis is not unique to ophthalmology, and other disciplines in biology were

[☆] Grant Information: Ophthalmic Research Institute of Australia, NHMRC Biomedical Postgraduate Scholarship (PGS), Pfizer Australia Senior Research Fellowship (DAM).

* Corresponding author to: David Mackey, Centre for Eye Research Australia, Royal Victorian Eye and Ear Hospital, 32 Gisborne St, East Melbourne, Victoria 3002, Australia. Tel.: +61 3 9929 8375; fax: +61 396623859.

E-mail address: d.mackey@utas.edu.au (D.A. Mackey).

facing similar restrictions, until the advent of geometric morphometric techniques that led to a “revolution” in biological shape analysis (Rohlf, 1990; Bookstein, 1997; Adams et al., 2004; Zelditch et al., 2004; Slice, 2007). Geometric morphometrics provides a statistical framework that retains the spatial relationships among regions within anatomical structures, providing a complete, uniform foundation for the quantitative shape analysis of biological forms. The discipline of geometric morphometrics is built upon methods that can be broadly classified as landmark-based, outline-based and a composite semi-landmark-based technique. A comprehensive review of the discipline and the individual techniques comprising it can be found elsewhere (Adams et al., 2004; Zelditch et al., 2004; Sanfilippo et al., 2009).

The purpose of the present study was to explore the application of geometric morphometrics methods in assessing optic disc morphology. Specifically, our goal was to determine the optimal geometric morphometric method for describing and quantifying optic cup shape of a sample of normal and glaucomatous optic discs. We used linear discriminant analysis (LDA) (Neff and Marcus, 1980), a widely used statistical technique that suits a variety of data including those from traditional and geometric morphometrics (Klingenberg and Monteiro, 2005). LDA has proven useful for elucidating characteristics of shape that help discriminate between groups in many biological and medical disciplines, from taxonomy and phylogenetics to forensics and biomedicine (Díaz-Flores Estévez et al., 2004; Cardini and Elton, 2008; Franklin et al., 2008; Bennett et al., 2009; Stefánsson et al., 2009). Low rates of misclassification provide empirical evidence that diagnostic shape features are extracted by the linear discriminant function (LDF) in allocating specimens to their true population.

2. Materials and methods

Optic disc data used in this study were obtained from simultaneous optic disc stereophotographs of Caucasian subjects recruited as part of the Glaucoma Inheritance Study of Tasmania (GIST) and the Twin Eye Study Tasmania (TEST). Photos were taken with a Nidek 3-Dx camera (Nidek, Gamagori, Japan) and the processed color 35 mm slides digitized at a high resolution (2102×1435 pixels, 2900 ppi, 36 bit colour) using a Nikon CoolScan IV ED slide scanner (Nikon Corp., Tokyo, Japan). An experienced observer masked to the subjects' disease status used custom software with a z-screen (Real D Corp. Beverly Hills, CA) to delineate the optic cup and quantify optic disc parameters following the technique described in detail elsewhere. (Sheen et al., 2004; Morgan et al., 2005a,b) Briefly, a standard CRT monitor and overlying high speed-modulating panel is used to provide flicker-free stereoscopy of component images of a stereo pair alternatively displayed at 60 Hz, when viewed with polarized glasses. The observer can adjust cursor depth to coincide with Elschnig's rim, such that the neuroretinal rim (NRR) and inner margin of the optic disc can be outlined at the depth of the scleral plane. Scaled estimates of optic disc parameters corrected for image magnification are calculated from keratometry readings, refraction and camera specifications using standard methods. Resultant measures of optic cup, disc and NRR areas, as well as outline coordinate data in 5° segments (72 points total) are produced.

For the study 160 subjects who had previous optic disc quantification using the z-screen were identified. Stereoscopic optic disc photos from 80 case subjects (GIST) with established glaucoma and an equal number of phenotypically normal control subjects (TEST) were analyzed. Left eye images were selected for analysis. Case subjects were required to have a borderline or abnormal Glaucoma Hemifield Test (Humphrey Field Analyzer II; Carl Zeiss Meditec, Inc., Dublin, CA) due to glaucomatous optic neuropathy (cup-disc

ratio ≥ 0.7 ; inter-eye disparity in cup-to-disc ratio ≥ 0.2 ; or focal NRR rim notching) in their left eye. This study was approved by the relevant ethics committees of the Royal Victorian Eye and Ear Hospital and the Royal Hobart Hospital and was conducted in accordance with the Declaration of Helsinki.

3. Optic cup shape quantification and discrimination

To determine the optimal geometric morphometric technique for measurement of optic cup shape, three methods of shape analysis relevant to biological forms without definable landmarks were assessed: Elliptic Fourier analysis (EFA), sliding semi-landmark analysis (SSLA) with a minimum bending energy (BE) criterion and SSLA with a minimum Procrustes distance (PD) criterion. These methods have been reviewed in detail elsewhere (Sanfilippo et al., 2009), and are described in Appendix 1.

EFA and SSLA were performed on the outline data to eliminate variation unrelated to shape (i.e. location, size and orientation) and yield a series of shape variables for each technique, which can be summarized using principal component analysis (see below), used to test group shape differences and analyzed to estimate the accuracy of discrimination. The effect of including a variable number of coordinate points in subsequent shape analyses was also explored. Of the original 72 (5° apart) outline coordinates generated by the Cardiff software, points were progressively removed to leave: 36 points (10° apart), 24 points (15° apart), 18 points (20° apart), and 12 points (30° apart).

3.1. Comparison with 'traditional measures' of optic cup shape

In order to assess how well EFA and SSLA perform in differentiating normal and glaucoma optic cup shapes, it is important to compare these techniques with those that are conventionally used for this purpose. For each subject, five aspects of optic cup shape were measured from the outline data: vertical-to-horizontal diameter ratio and maximal-to-minimal diameter ratio (these two as measures of cup ovality), angle between maximal diameter and horizontal, form-factor and inferior-to-temporal NRR width ratio. Differences in mean values between normal and glaucoma groups were tested for statistical significance and the discriminatory power of these variables assessed.

3.2. Discriminant analysis

We calculated LDFs from the results of all geometric morphometric techniques and traditional measures and assessed each method's efficacy in classifying normal and glaucomatous cups. An LDF represents a linear combination of the set of original variables, built in order to maximize between relative to within sample variance. In this setting, each function was derived from variables that describe shape features of optic cup shape. To avoid the problem of inflated classification rates that occur when discriminant functions are evaluated on the same data they were derived from, we employed a leave-one-out cross-validation approach. This entails leaving each observation out in turn and computing a discriminant score for the excluded observation based on the discriminant function derived from the remaining $N - 1$ observations. The sensitivity and specificity of each technique was then calculated as the proportion of correctly classified glaucomatous and normal optic cups, respectively. We then computed a receiver operator characteristic (ROC) curve from the discriminant analysis posterior probabilities for each geometric morphometric technique examined and determined the area under the curve (AUC). The ROC curve is a graphical representation of the true positive rate (sensitivity) compared with the false positive rate ($1 - \text{specificity}$) for

a dichotomous classification system as its discrimination threshold is varied (Hanley and McNeil, 1982). The AUC values were computed from ROC curves derived for each technique with higher values demonstrating superior discrimination ability (an AUC of 1.0 represents perfect discrimination and an AUC of 0.5, chance discrimination).

3.3. Stratification by glaucoma severity

The technique with the highest AUC was subsequently selected to investigate the discrimination efficacy of optic cup shape information in the glaucoma group stratified by severity. Visual field plots were examined and an AGIS (Advanced Glaucoma Intervention Study) score determined for each individual (AGIS Investigators, 1994). The AGIS score is based on both the number and depth of adjacent test locations and provides a method of glaucoma severity staging from visual field data, such that: Score 0 – normal visual field; Score 1–5 mild damage; Score 6–11 moderate damage; Score 12–17 severe damage; Score 18–20 end-stage. For the purposes of this study we combined the mild and moderate and severe and end-stage groups to provide three categories of visual field damage: Pre-perimetric, Moderate and Severe.

3.4. Classification by trained graders

The quantitative evaluation of the performance of the morphometric techniques for distinguishing between normal and glaucomatous cups was compared with that of two graders masked to disease status. Both graders are ophthalmologists trained to fellowship standard in glaucoma and experienced in biomicroscopic examination of the optic disc. The outlines of all optic cups were reconstructed from the coordinate data generated by the Cardiff software and scaled to the same size. The observers undertook two grading tasks. In the first session, an image of the outline of each optic cup was randomly presented on a computer screen with no additional subject or photographic information and the grader was asked to indicate disease status according to an ordinal ranking of five confidence levels: definitely normal, probably normal, undecided, probably glaucoma, and definitely glaucoma (Fig. 1a). Confidence levels were weighted on a scale of 0 (definitely normal) to 1 (definitely glaucoma) reflecting the certainty of the grader of a correct classification and used to construct a combined ROC curve for both graders. A training set of five normal and five glaucoma cup outlines were viewed prior to commencing the test. One week later each grader was asked to classify the stereophotos from which the outline data were derived,

and asked to indicate disease status using the same schema as above (Fig. 1b). Stereoscopic visualization of the optic disc was achieved by viewing each photo through a handheld stereo-viewer. Within and between grader repeatability was assessed by re-grading 20 random outline images for each task and calculating an intra-class correlation coefficient (ICC) (Shrout and Fleiss, 1979).

4. Morphometric and statistical analysis

Univariate group differences were tested using t-tests for each traditional cup shape variable one at a time. For shape, principal component analysis (PCA) was used as a dimensionality reduction technique for the variables derived from EFA and SSLA (EFA coefficients, SSLA Procrustes shape coordinates) prior to subsequent analyses (see also Appendix). Variation on the main PC axes was illustrated using bivariate scatterplots. Group differences were tested using the multivariate extension of a permutation test for sample mean differences (Manly, 1997), based on mean shape distances. Permutation tests tend to be more conservative than their parametric counterparts, but make fewer assumptions and can also be used when sample size is small relative to the number of variables, a common occurrence in shape analysis (Bookstein et al., 1999).

Excluding area and cup-to-disc size measurements, the relative importance of each of the traditional cup shape variables was assessed by stepwise LDA. LDA was performed on all principal component (PC) axes with non-zero eigenvalues derived from the PCA.

Statistical analyses were conducted using SPSS (v17.0; SPSS, Chicago, IL). EFA was performed using Morpheus (Slice, 1999). The EFA coefficients were imported into SPSS for PCA and LDA, and in NTSYSpc (Rohlf, 2006) for permutation tests. Procrustes superimposition and sliding of semi-landmark points (minimum BE and minimum PD criteria) was initially conducted in tpsRelw (Rohlf, 2008) and the aligned data then imported into MorphoJ (Klingenberg, 2008) for PCA, LDA and permutation tests of shape differences. ROC curves were generated in SPSS.

5. Results

5.1. Patterns in optic cup shape variation

Based on the discriminant analyses and ROC curve results (see below), a SSLA approach (18 points, min BE criterion) was selected as the one with greatest accuracy and used to quantify and illustrate the variation in optic cup shape in the dataset. Over

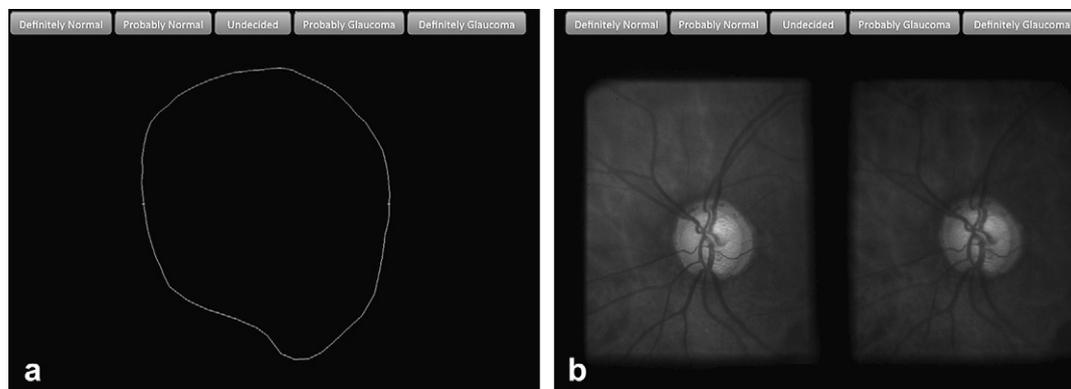


Fig. 1. a (left), b (right) – Digital images of optic cup shape: outline only (a) and optic disc stereophotographs (b). Two graders were asked to classify the optic discs as normal or glaucomatous. The series of outlines were presented at the first grading session, followed by the stereophotos as the second session one week later.

half the variation in the shape variables was explained by the first PC and approximately 90% by the first four PCs. Inspection of the individual optic cup scores in the PCA scatterplots for the first four dimensions of shape variation demonstrate several patterns of interest (Fig. 2). The remaining PCs and the visualization of their respective shape changes did not reveal any evident biologically relevant feature. Ordination of the data along the first two PCs shows a weak clustering effect with more normal optic cups having a positive score and most glaucomatous optic cups a negative score in these two dimensions (Fig. 2a).

The third PC accounts for approximately 11% of the variance in shape and describes a localized variation that produces nasal contraction/temporal expansion of the optic cup outline inferiorly when the PC score is positive. As with the first two PCs, a weak stratification of points is observable in the PCA scatterplot along

this axis (Fig. 2b). It is difficult to identify a pattern of clustering along the 4th PC. The patterns of optic cup shape change associated with each PC are shown in Fig. 3.

5.2. Optic cup shape discrimination

Group differences were tested for all sets of shape variables. Mean values for each of the traditional measures of optic cup shape are provided in Table 1. Using geometric morphometric techniques, optic cup shape was found to differ significantly between normal and glaucoma groups ($p \leq 0.0005$, 10,000 permutations including observed) in all EFA and SSLA analyses. The mean shapes for each group are displayed in Fig. 4 and the deformation associated with the mapping of the glaucomatous configuration onto the normal

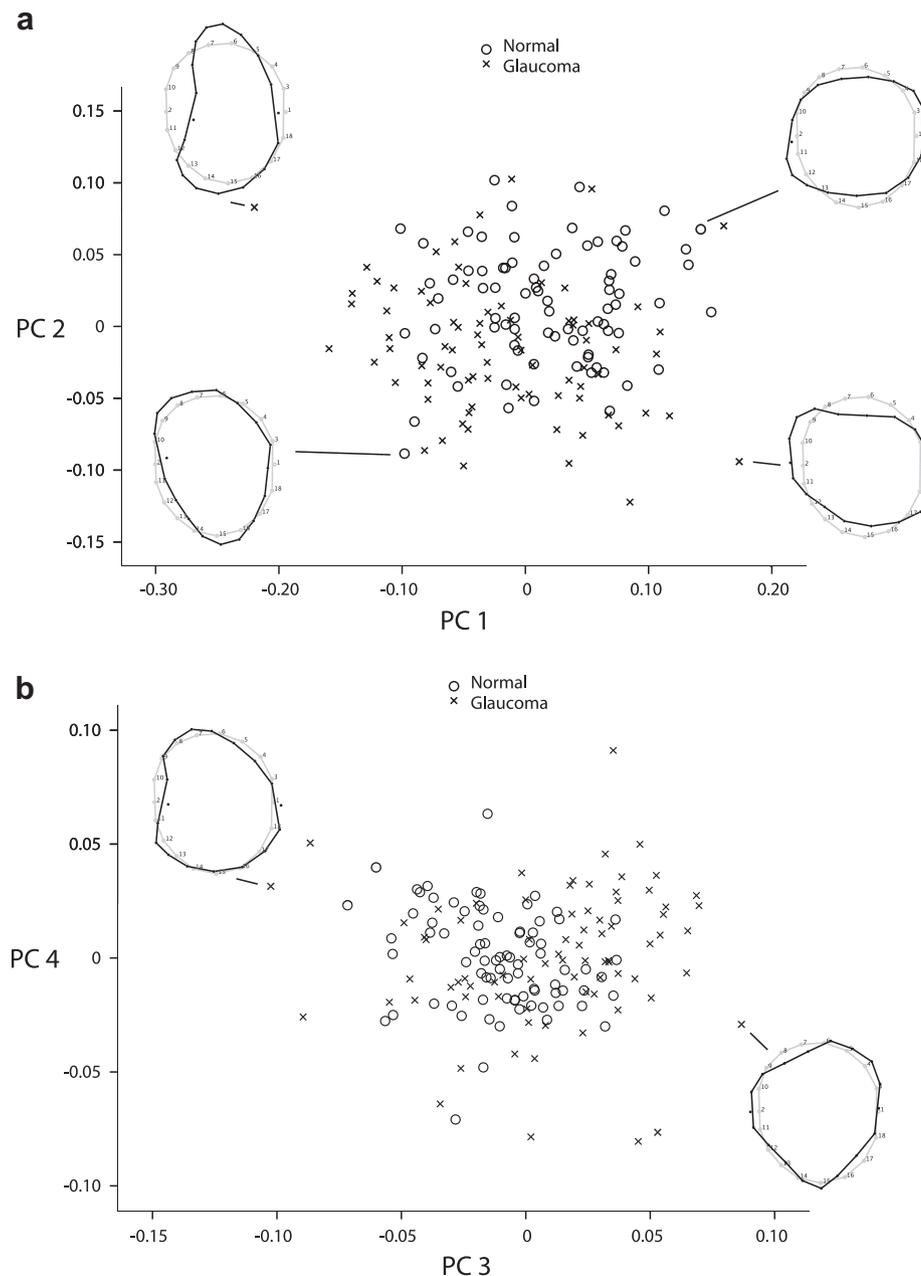


Fig. 2. a (top), b (bottom) – Plots of the first four dimensions of a PCA of optic cup shape; PC1 vs PC2 (a), PC3 vs PC4 (b). Example cases are shown as outlines to illustrate the shape variations (size variation removed). The light outline represents the mean shape and the dark outline the optic cup for the indicated subject.

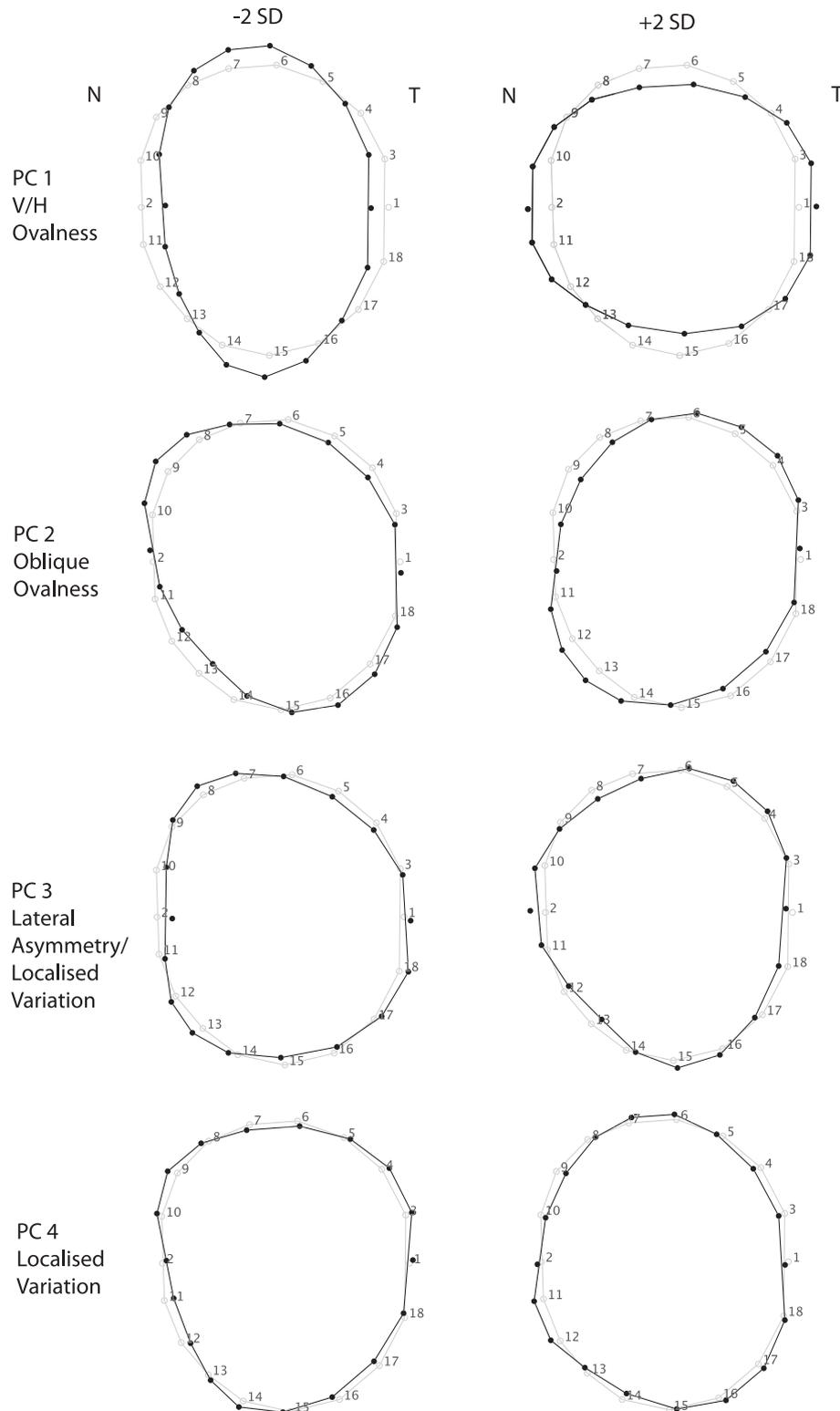


Fig. 3. Visualization of shape variation on the first four PCs relative to mean shape (light outline) of the optic cup sample. The dark outline represents the mean shape ± 2 SD from the mean for that PC.

one represented graphically using deformation grids and displacement vectors.

Table 2 shows the effectiveness of each technique in differentiating normal from glaucomatous optic cups based on shape information alone. The cross-validated classification rates (sensitivity and specificity) were determined from the calculated discriminant

functions for each technique. The higher the classification rate (%), the better the technique is at allocating each optic cup to its a priori determined disease status group.

SSLA optimized with a bending energy criterion and, using 18 of the original 72 points, had the best sensitivity (83%) and AUC value (0.91) (Fig. 5). EFA (72 points) performed similarly well (74%, 0.89).

Table 1

Optic disc characteristics and traditional optic cup shape measures for normal and glaucoma subjects (mean \pm SD). The form-factor measures outline undulation and varies from one for a circle to zero as the structure differentiates from a circle. The neuroretinal rim width ratio represents the ratio of the inferior to temporal NRR dimensions. The variables in bold were used to generate the ROC curve for traditional measures described in Table 2.

Optic Disc Characteristics	Glaucoma	Normal	P value
Sample size	80	80	
Age at evaluation	70.0 \pm 10.0	55.5 \pm 10.5	<0.0001
Optic disc area (mm ²)	2.14 \pm 0.47	2.02 \pm 0.54	0.86
Neuro-retinal rim area (mm ²)	1.16 \pm 0.36	1.57 \pm 0.43	<0.0001
Cup/disc area ratio	0.45 \pm 0.16	0.24 \pm 0.20	<0.0001
<i>Traditional optic cup shape measures utilized in LDA</i>			
Cup Ovality (vertical/horizontal diameter)	1.16 \pm 0.17	1.07 \pm 0.12	0.0001
Cup Ovality (maximal/minimal diameter)	1.33 \pm 0.17	1.20 \pm 0.10	<0.0001
Angle between maximal diameter and horizontal (°)	103.5 \pm 29.6	99.4 \pm 40.0	0.54
Form-Factor (cup)	0.93 \pm 0.07	0.86 \pm 0.14	0.0003
Neuroretinal rim width ratio	1.39 \pm 0.62	1.34 \pm 0.37	0.44

When all points were utilized in a SSLA based on a minimal Procrustes distance criterion, the sensitivity and AUC were poorest (55%, 0.70). Of the traditional shape-based variables, ovality (maximal-to-minimal diameter) and form-factor were the only significant variables in the stepwise LDA model. The discriminant function calculated from these two measures provided comparable classification rates and AUC (0.86) to the geometric methods, although the capacity to classify normal optic cups correctly (specificity) was somewhat reduced (75%).

Classification of glaucoma severity by AGIS score revealed 20% had normal visual fields, 30% mild damage, 17.5% moderate damage, 22.5% severe damage and 10% end-stage glaucoma. Optic cup shape did not significantly differ between each of the three AGIS-based stages we employed in this study. LDA classification rates between pre-perimetric, moderate and severe groups are shown in Table 3.

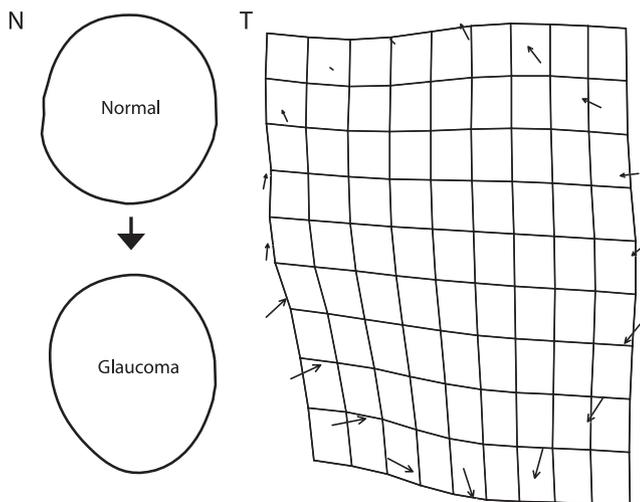


Fig. 4. Thin-plate spline (TPS) deformation grids demonstrating the difference in mean shape between normal and glaucomatous optic cups. The TPS expresses differences in shapes (normalized for size) as a continuous deformation derived from a smoothing function applied to the landmark configuration (Bookstein, 1989). There is greater warping of the grid inferiorly reflecting a relative vertical elongation in the average glaucoma shape. Additionally, the cup expansion tends to be infero-temporally directed and may be representative of the pattern of NRR excavation that occurs in glaucoma. The displacement vectors represent the overall pattern of shape change for the deformation of normal to glaucomatous cupping in this example and should not be interpreted as movements of specific landmarks.

For the graders, the sensitivities and specificities for detecting glaucoma exclusively from optic cup shape (outline) information were 80% and 56% (Grader 1) and 55% and 84% (Grader 2), respectively. When the same classification task was repeated with the optic disc stereophotos, the overall proportion of correct assignments increased, with sensitivities and specificities of 76% and 84% (Grader 1) and 84% and 88% (Grader 2), respectively. The AUCs calculated from the combined grader data were 0.74 (outline) and 0.88 (photos). Intra-grader repeatability was 0.72 (0.54–0.84, $p < 0.0001$) (Grader 1) and 0.69 (0.48–0.82, $p < 0.0001$) (Grader 2) and the agreement between graders 0.64 (0.49–0.75, $p < 0.0001$).

6. Discussion

This study evaluated modern geometric morphometric approaches for the measurement and description of optic cup shape and compared the results to traditional univariate measures. Although both approaches performed well and almost all sets of predictors showed highly significant differences in optic cup shape between normal and glaucomatous individuals, the geometric morphometric techniques were better at discriminating optic cups, and provided a more comprehensive description of shape variation in the datasets than that produced by the traditional measures.

Vertical elongation and supero-temporal/infero-temporal expansion of the optic cup contour were reported by both graders as the most important aspects of cup shape variation in assigning disease status based solely on outline data. The primary shape characteristics of the optic cups in this study, as revealed by PCA, are displayed graphically in Fig. 3. The shape changes associated with the first two PCs may be interpreted as ovalness of the optic cup along the cardinal and oblique axes. More negative PC scores correspond to a vertically elongated cup on the first PC axis and an infero-temporally elongated cup on the second PC axis. Both of these morphological cup patterns are characteristic of glaucoma. Given that the dataset is composed of left optic cups and more glaucomatous specimens scored positively along the third PC axis, the observation concurs with the pathognomonic importance of the infero-temporal NRR in glaucoma.

Optic cup and disc shape have traditionally been evaluated in terms of ovality and form-factor, with the first measure describing elongation and the second undulation of the outline. Indeed, cup ovality (maximal-to-minimal diameter) and form-factor were selected by the stepwise LDA as the best combination of traditional variables, among those not involved in the original diagnosis, to discriminate normal and glaucomatous optic cups. Elongation of the optic cup in this sample of optic discs was greater in the glaucoma group for the two ovality criteria assessed, which corresponds to a vertical deepening concordant with the pattern of NRR excavation described in the literature (Anderson and Drance, 1995). Interestingly, the NRR width ratio did not reflect this as might have been expected. However knowledge of the rim configuration and optic disc dimensions is necessary to account for this finding. Although unreported in previous studies (Jonas et al., 1988c; Jonas and Papastathopoulos, 1996), our data suggest a possible explanation: that the vertical dimension of the optic disc is also greater in eyes affected by glaucoma, an aspect that will be investigated in future studies on the association between shape change and the development of glaucoma.

Discriminant functions were computed for the group of traditional univariate measures of cup shape and for the scores of the PCA on the shape variables of each geometric morphometric technique. The resultant test performance parameters (sensitivity, specificity, AUC) enabled a direct comparison between each technique's ability to assess cup shape. The ability to classify a cup as glaucomatous was generally high (>70%) for all methods, but the

Table 2

Classification rates and area under the ROC Curve calculated from discriminant functions for each geometric morphometric technique. Abbreviations: CI, confidence interval; EFA, elliptic Fourier analysis; SSLA, sliding semi-landmark analysis; min BE, minimum bending energy; min PD, minimum Procrustes distance.

Morphometric technique/Grader	Number of points used	Cross-Validation Classification Rate (%)		Area under the ROC curve (95% CI)	P value
		Glaucoma (Sensitivity)	Normal (Specificity)		
Traditional Measures	NA	76	75	0.86 (0.80–0.92)	0.0001
EFA	18	74	86	0.88 (0.83–0.92)	0.0001
EFA	36	74	86	0.88 (0.83–0.92)	0.0001
EFA	72	74	88	0.89 (0.83–0.93)	0.0001
SSLA (min BE)	12	78	81	0.85 (0.79–0.90)	0.0001
SSLA (min BE)	18	83	86	0.91 (0.85–0.95)	0.0001
SSLA (min BE)	24	70	88	0.87 (0.81–0.92)	0.0001
SSLA (min BE)	36	66	89	0.82 (0.75–0.88)	0.0001
SSLA (min BE)	72	56	83	0.76 (0.66–0.83)	0.0001
SSLA (min PD)	12	76	84	0.88 (0.81–0.92)	0.0001
SSLA (min PD)	18	78	83	0.86 (0.79–0.91)	0.0001
SSLA (min PD)	24	70	83	0.85 (0.79–0.90)	0.0001
SSLA (min PD)	36	63	85	0.79 (0.72–0.86)	0.0001
SSLA (min PD)	72	55	87	0.70 (0.63–0.79)	0.0001
Combined Grader Results (cup outline only)	NA	68	68	0.74 (0.69–0.79)	0.0001
Combined Grader Results (stereophoto)	NA	80	86	0.88 (0.84–0.91)	0.0001

assignment rates were quite variable. Traditional shape variables compared well with geometric morphometric techniques on sensitivity, with 76% for the traditional measures and between 55 and 83% for geometric morphometric techniques (SSLA min PD with 72 points and SSLA min BE with 18 points, respectively). On specificity, traditional techniques were not as good as geometric morphometric techniques, with 75% for the traditional measures and between 81 and 89% for the geometric morphometric techniques (SSLA min BE with 12 and 72 points, respectively). The consistent intra-technique disparity favoring specificity over sensitivity may be due to the greater variability in cup shape of glaucoma compared with normal optic discs. As a measure of overall performance, the AUC was similar among all techniques examined, with less than 0.06 separating the best performing technique from EFA and the traditional measures.

For comparison, two graders used their clinical expertise in diagnosing glaucoma to assign disease status to each optic disc

from photos and cup outlines. The average sensitivity and specificity for both graders in this study (80%, 86%) were similar to results found by Morgan and colleagues upon stereoscopic classification of the optic disc (83%, 85%) (Morgan et al., 2005b). As a gold standard technique, the AUC calculated from stereoscopic assessment of optic disc photos was high (0.88), indicating low rates of misclassification. One geometric morphometric technique, SSLA min BE with 18 points, compared favorably with the graders, having slightly better sensitivity (83 vs. 80%), the same specificity (86%), and slightly higher AUC (0.91 vs. 0.88). This suggests that cup shape characteristics as assessed by geometric morphometric techniques may be effective in glaucoma diagnosis. Hewitt et al. showed that the shape of the optic cup and disc were useful morphological features of the optic disc in determination of twin zygosity (Hewitt et al., 2007). As expected, the performance of each grader decreased when all optic disc information except that available in the cup outline was removed. Interestingly, the graders differed in their diagnostic approach for this difficult task, with Grader 1 biased towards not missing disease (higher sensitivity) and Grader 2 not overcalling disease (greater specificity).

While the shape of the optic cup is an effective discriminator of normal and glaucomatous optic discs, it performs less well in correctly staging individuals with established disease in these data (i.e., when glaucomatous discs only are analyzed). In most cases, the cross-validated classification rates of individuals staged by severity of glaucoma are no better than chance. This may be related to low statistical power in smaller samples. It is also likely that the shape features that are relevant in distinguishing between normal and glaucomatous individuals may already be present even in pre-perimetric glaucoma. This is difficult to ascertain in our study because even though the AGIS classification stratifies glaucoma

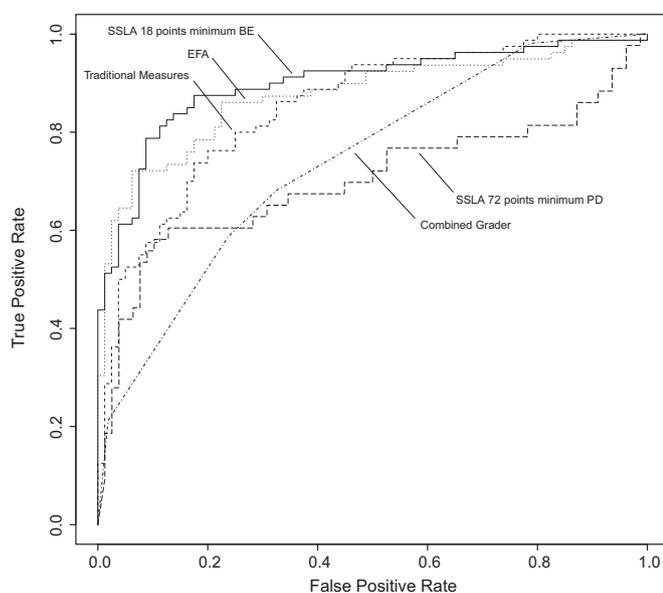


Fig. 5. A comparison of ROC curves for the best and poorest performing geometric morphometric techniques in discriminating normal and glaucoma cups. The ROC curves for EFA, the traditional shape measures and the combined graders (cup outline only) performance have also been included. All other techniques fell within this range.

Table 3

Classification rates for the SSLA (minimum bending energy criterion – 18 points) technique stratified by visual field severity.

AGIS groups tested	Cross-validation classification rate (%)
Pre-perimetric	50
Moderate	61
Moderate	50
Severe	46
Pre-perimetric	56
Severe	54

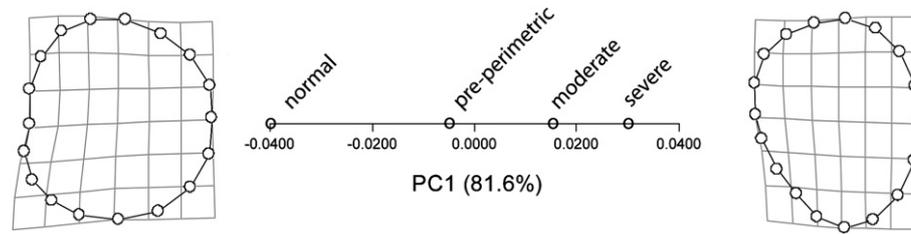


Fig. 6. PCA of normal and AGIS stages mean shapes: PC1 (81.6% of total variance) and extremes of variation along the vector shown using TPS deformation grids (magnified two times).

severity blind to disc appearance, optic disc characteristics were considered in the original diagnosis. Indeed, a potential limitation of our study relates to a risk of circular reasoning in establishing the importance of shape as a diagnostic optic disc parameter given that shape features may have been initially evaluated. An ideal study design in this case would be to ascertain subjects with diagnosed glaucoma masked to optic disc status (a difficult task as this is the primary pillar of glaucoma diagnosis). As proof of principle, our glaucoma cohort consists of subjects with established disease. Nonetheless it is worth noting that despite the large overlap in optic cup shape variation of glaucomatous individuals leading to poor discrimination accuracy within this group, a PCA of average shapes for the three stages of glaucoma and the normal group (Fig. 6) shows a main axis of change ($\sim 82\%$ of total variance) that perfectly aligns with the direction set by the severity of the disease (i.e., Normal \rightarrow Pre-perimetric \rightarrow Moderate \rightarrow Severe) as assessed by AGIS score. A further limitation of our study is the non-availability of optic cup shape three-dimensional data. In our case this is a methodological limitation at the time of data acquisition, and the extension to 3-D analysis in geometric morphometrics is straightforward using the same principles. However, even without using this information, we were able to capture diagnostically relevant shape features. Adaptation of the methodology to enable 3-D analysis will be necessary to reveal the complete significance of shape as a classifier variable. Lastly, with improvements in automated segmentation algorithms, future work to demarcate the cup outline should ideally be conducted without the potential for observer bias.

Linear discriminant analysis has become increasingly common for methodological and instrument validation in ophthalmology. The advent of computerized technology for imaging the optic disc has facilitated the rapid acquisition of rich morphological data and driven the desire to develop techniques for improved glaucoma detection by automated means, such as the Glaucoma Probability Score classification algorithm of the Heidelberg Retina Tomograph (HRT-3, Heidelberg Instruments, Heidelberg, Germany). In contrast to more complex machine learning or neural network classifiers; however, the examination of linear discriminant models is relatively simple. Most generic statistical packages can perform discriminant analysis and the underlying mathematical principles are not conceptually difficult.

Several groups have investigated the efficacy of discriminant functions based on HRT-measured cup characteristics. Pablo and colleagues reported that the diagnostic accuracy of the HRT-3 algorithms was comparable with that obtained in stereophoto evaluation by an experienced glaucoma specialist (Pablo et al., 2009). They developed an LDF for HRT-3 parameters that provided a better sensitivity-specificity balance compared to the standard instrument algorithms (Ferrerres et al., 2008). Iester and colleagues used ROC curves to demonstrate that the 'cup shape measure' was the best HRT-measured variable for differentiating normal and glaucomatous eyes (cup shape in this context refers to

the shape of the frequency distribution of depth values within the optic cup) (Iester et al., 1997). Bathija et al. advocated the use of a combination of cup shape measure, rim area, retinal nerve fibre layer (RNFL) thickness and height variation contour for establishing disease status (Bathija et al., 1998).

Gunvant and coworkers investigated the performance of Fourier analysis applied to RNFL thickness measurements obtained from optical coherence tomography (OCT) in classifying normal and glaucoma eyes (Gunvant et al., 2007). Discriminant functions derived from Fourier coefficients representative of the characteristic 'double hump' shape of the RNFL TSNIT (temporal, superior, nasal, inferior, temporal) curve were superior to the standard OCT thickness output measures.

An interesting finding in this study was that increasing the number of points used to describe the outlines did not always lead to better results. EFA was relatively insensitive to the number of points, which is not unexpected given the simplicity of the shapes analyzed and that the Fourier function is summed over all points. Besides, by including only coefficients from the first eight harmonics, small-scale features, which may have increased noise, were filtered out in the analysis. SSLA, however, produced best results with 18 points, which corresponds to points spaced 20° apart. A possible explanation for this finding is that increased detail in outline shape captured with additional points contributes as much to the variability in shape as does noise. We are not the first to observe such effect. In a study of human craniofacial and dental variation using SSLA, Perez et al. similarly found an improvement in cross-validated classification rates when the number of outline points used was reduced from 82 to 44 (min BE criterion 67.86%, 73.21%, respectively) (Perez et al., 2006).

To crudely assess our impression that most diagnostic shape variation concerns large-scale features that are efficiently described using relatively few semi-landmarks, we performed an LDA on the same data utilizing only the uniform and a subset of the remaining non-uniform shape variables (partial warps). Partial warps and uniform components are descriptors of a particular shape change that are derived to compute deformation grids. As such, they are biologically arbitrary and as a general rule should not be analyzed individually or as a subset (Rohlf, 1998). However, as the uniform components and higher-order partial warps correspond to large-scale shape changes, and lower-order partial warps to small scale-shape variation, we used them in an ad-hoc fashion to investigate the potential reasons for the poorer accuracy associated with an increased number of outline points. When the uniform and 18 highest-order partial warps were considered, cross-validated discriminant scores resulted in a 75.0% sensitivity and 81.3% specificity. In contrast, an LDA on the remaining 120 partial warps produced poor results in term of accuracy (51.3% sensitivity and 78.8% specificity), similar to those from the whole set of shape variables (Table 2). Indeed, as we suspected, large-scale features, mostly detected by uniform components and high-order partial warps, increased the classification accuracy and especially the

sensitivity to a remarkable degree. The effect of 'noise' in small-scale outline detail was potentially increased in the glaucomatous group given the greater variability in shape and larger cup size, thus highlighting the importance of accurate curve fitting by the digital planimetry software to the observer-defined cup outline. Additionally, fewer of the glaucomatous outlines were required to be smoothed (least-squares algorithm) prior to analysis so as to reduce irregular variation induced by the software contour fitting of very small cups. The differences introduced by smoothing were not significant compared to individual and group variation (results not shown). However, the larger variance observed in the glaucomatous sample, regardless of smoothing, became even greater relative to the normal sample when small-scale detail of little diagnostic relevance was detected in the SSLA analysis using all 72 points. This suggests that several solutions are available to reduce noise in data such as these. Decreasing the number of outline points and performing a SSLA with a minimum bending energy criterion seems parsimonious and effective, as indicated by unchanged high sensitivity (83%) and specificity (86%) even when no smoothing was applied (results not shown).

The main aim of this study was to assess the relative discriminatory power of different methods to describe the shape of the optic disc. In this respect, our analysis has shown that a geometric morphometric approach to the discrimination of optic cup shape between normal and glaucomatous eyes is superior to that provided by traditional single parameter shape measures, and comparable with that of two graders using stereophotos. The additional value in a geometric methodology is its ability to measure and map shape variation onto the study structure to quantitatively describe the resultant patterns graphically. The application of geometric morphometric principles within the field of ophthalmology has been limited to date, and it is hoped that this study has demonstrated the potential power and utility of such methods in evaluating shape variation, providing effective visualization, and accurately discriminating between groups. Such analytical techniques could be incorporated into future automated optic disc screening modalities.

Acknowledgments

The authors wish to thank Sandra Staffieri for help with retrieval of visual field data.

Appendix 1. Description of geometric morphometric techniques utilized

Elliptic Fourier analysis (EFA)

EFA describes shape mathematically by transforming coordinate data regarding contour information into Fourier coefficients. It achieves this by using a parametric representation of an outline in which the x and y co-ordinates of points are treated separately as functions of cumulative distance along the outline from an arbitrary starting point (Cardini and Slice, 2004). It then becomes possible to express x and y co-ordinates (for given values of t) along an outline as an infinite sum of sine and cosine terms (harmonics). Plotting the separate harmonics produces a set of orthogonally related ellipses that can be summed to reconstruct the original outline. Additional harmonics increase resolution, improving the ability to detect more refined and localized shape differences. However, to constrain the dataset for efficient statistical analysis, an optimal number of harmonics is usually selected to exclude coefficients containing negligible shape information. Reconstructed outlines can be assessed for accuracy by calculating an 'error' or residual value that represents the difference between the original

data and the predicted values from EFA. Fourier coefficients were computed and outlines reconstructed for each subject using a variable number of harmonics (1–15,20) to provide 72 evenly spaced points (as originally generated by the Cardiff software). The sum of squared distances between each of the original and corresponding predicted points was compared with the number of harmonics used for the reconstruction. The greatest decrease in error was noted with the inclusion of the first four harmonics and conservatively remained unchanged after the inclusion of eight. Therefore, the coefficients of the first eight harmonics were used in further analyses of optic cup shape.

Normalization of the data to remove location, scale and orientation effects is an important issue that needs to be addressed before meaningful shape comparisons can be made. To make the coefficients invariant to non-shape parameters, Kuhl and Giardina based their procedure on the best fitting ellipse represented by the coefficients of the first harmonic (Kuhl, 1982). Once the coefficients have been normalized, they can be used as shape variables in multivariate analyses.

Even though an individual outline may have an irregular and characteristic shape, the first harmonic ellipse may be almost circular, making it difficult to identify the major axes, which in turn prevents standardizing the sample orientation. To avoid this problem we standardized orientation using the two points at the intersection of the outline with a horizontal (nasal-temporal) line passing through the outline centroid. These two coordinates were then used as 'landmarks' in a baseline Bookstein superimposition to normalize each configuration prior to an EFA to characterize the outlines themselves.

Sliding semi-landmark analysis (SSLA)

The sliding semi-landmark methodology incorporates information on curves into the landmark formalism (Bookstein, 1997). Semi-landmarks are essentially a series of points located along a contour whose position is defined using non-biological criteria (equally spaced points, points halfway between the ends of a curve connecting two landmarks, etc.) (Zelditch et al., 2004). Even though semi-landmarks are not homologous across specimens, the outline curve or contour under examination should be. Thus, in order to generate meaningful shape coordinates, algorithms may be used to reduce the variance in semi-landmark coordinates that may be otherwise inflated due to the lack of a simple one to one inter-individual correspondence between these points.

Following standard Procrustes superimposition whereby specimen configurations are made invariant to position, size and orientation, semi-landmark points on each specimen can be permitted to 'slide' along the outline curve (Adams et al., 2004). Two sliding semi-landmark criteria have been developed to facilitate this objective: minimum Procrustes distance and minimum bending energy. With both techniques, semi-landmarks are slid in either direction along the curve, which is approximated by a line parallel to the segment connecting two adjacent points (Rohlf, 2008). Then, if the bending energy is minimized, the set of slid semi-landmarks will represent the smoothest deformation of the curve on the reference form to the corresponding curve on an individual specimen. Alternatively, if the Procrustes distance is minimized, the set of slid semi-landmarks will be at a minimum distance between the curve on the reference form and that of the individual. Once the semi-landmarks are optimally adjusted, they can be treated as shape variables in statistical analysis. The reduction of variance after sliding implies a loss of degrees of freedom beyond that customarily due to the Procrustes superimposition (i.e., 4 d.f. in a two-dimensional analysis). Performing a principal component analysis on the shape coordinates and

discarding principal components with zero eigenvalues can effectively remove redundancy in shape variables.

As with our technique for EFA, we determined nasal and temporal 'landmark' coordinates for each optic cup configuration and treated them as homologous landmarks to be Procrustes superimposed whilst all other points were slid to minimize either bending energy or Procrustes distance.

References

- Adams, D.C., Rohlf, F.J., Slice, D., 2004. Geometric morphometrics: ten years of progress following the 'revolution'. *Ital. J. Zool.* 71, 5–16.
- Anderson, D.R., Drance, S.M. (Eds.), 1995. *Optic Nerve in Glaucoma*. Kugler Publications, p. 107.
- Advanced Glaucoma Intervention Study. 2. Visual field test scoring and reliability. *Ophthalmology* 101, 1994, 1445–1455.
- Bathija, R., Zangwill, L., Berry, C.C., Sample, P.A., Weinreb, R.N., 1998. Detection of early glaucomatous structural damage with confocal scanning laser tomography. *J. Glaucoma* 7, 121–127.
- Bennett, M.R., Harris, J.W.K., Richmond, B.G., Braun, D.R., Mbua, E., Kiura, P., Olago, D., Kibunjia, M., Omuombo, C., Behrensmeier, A.K., Huddart, D., Gonzalez, S., 2009. Early hominin foot morphology based on 1.5-million-year-old footprints from Ileret, Kenya. *Science* 323, 1197.
- Bookstein, F.L., 1989. Principal warps: thin-plate splines and the decomposition of deformations. *IEEE Trans. Pattern Anal. Mach. Intell.* 11, 567–585.
- Bookstein, F.L., 1997. Landmark methods for forms without landmarks: morphometrics of group differences in outline shape. *Med. Image Anal.* 1, 225–243.
- Bookstein, F.L., Schäfer, K., Prossinger, H., Seidler, H., Fieder, M., Stringer, C., Weber, G.W., Arsuaga, J.L., Slice, D.E., Rohlf, F.J., Recheis, W., Mariam, A.J., Marcus, L.F., 1999. Comparing frontal cranial profiles in archaic and modern homo by morphometric analysis. *Anatomical Record* 257, 217–224.
- Cardini, A., Slice, D., 2004. Mandibular shape in the genus *Marmota*: a preliminary analysis using outlines. *Ital. J. Zool.* 71, 17–25.
- Cardini, A., Elton, S., 2008. Variation in guenon skulls (I): species divergence, ecological and genetic differences. *J. Hum. Evol.*
- Díaz-Flores Estévez, J.F., Díaz-Flores Estévez, F., Hernández Calzadilla, C., Rodríguez Rodríguez, E.M., Díaz Romero, C., Serra-Majem, L., 2004. Application of linear discriminant analysis to the biochemical and haematological differentiation of opiate addicts from healthy subjects: a case-control study. *Eur. J. Clin. Nutr.* 58, 449–455.
- Ferreras, A., Pablo, L.E., Larrosa, J.M., Polo, V., Pajarín, A.B., Honrubia, F.M., 2008. Discriminating between normal and glaucoma-damaged eyes with the Heidelberg Retina Tomograph 3. *Ophthalmology* 115, 775–781.
- Franklin, D., O'Higgins, P., Oxnard, C.E., Dadour, I., 2008. Discriminant function sexing of the mandible of indigenous South Africans. *Forensic Sci. Int.* 179, 84.e1–84.e5.
- Gloster, J., 1975. Vertical ovalness of glaucomatous cupping. *Br. J. Ophthalmol.* 59, 721–724.
- Gunvant, P., Zheng, Y., Essock, E.A., Parikh, R.S., Prabakaran, S., Babu, J.G., Shekar, G.C., Thomas, R., 2007. Application of shape-based analysis methods to OCT retinal nerve fiber layer data in glaucoma. *J. Glaucoma* 16, 543–548.
- Hanley, J.A., McNeil, B.J., 1982. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* 143, 29–36.
- Hewitt, A.W., Poulsen, J.L., Alward, W.L.M., Bennett, S.L., Budde, W.M., Cooper, R.L., Craig, J.E., Fingert, J.H., Foster, P.J., Garway-Heath, D.F., Green, C.M., Hammond, C.J., Hayreh, S.S., Jonas, J.B., Kaufman, P.L., Miller, N.R., Morgan, W.H., Newman, N.J., Quigley, H.A., Samples, J.R., Spaeth, G.L., Pesudovs, K., Mackey, D.A., 2007. Heritable features of the optic disc: a novel twin method for determining genetic significance. *Invest. Ophthalmol. Vis. Sci.* 48, 2469–2475.
- Iester, M., Mikelberg, F.S., Swindale, N.V., Drance, S.M., 1997. ROC analysis of Heidelberg retina tomograph optic disc shape measures in glaucoma. *Can. J. Ophthalmol.* 32, 382–388.
- Jonas, J.B., Gusek, G.C., Guggenmoos-Holzmann, I., Naumann, G.O., 1988a. Size of the optic nerve scleral canal and comparison with intravitreal determination of optic disc dimensions. *Graefes Arch. Clin. Exp. Ophthalmol.* 226, 213–215.
- Jonas, J.B., Gusek, G.C., Guggenmoos-Holzmann, I., Naumann, G.O., 1988b. Variability of the real dimensions of normal human optic discs. *Graefes Arch. Clin. Exp. Ophthalmol.* 226, 332–336.
- Jonas, J.B., Gusek, G.C., Naumann, G.O., 1988c. Optic disc morphometry in chronic primary open-angle glaucoma. I. Morphometric intrapapillary characteristics. *Graefes Arch. Clin. Exp. Ophthalmol.* 226, 522–530.
- Jonas, J.B., Gusek, G.C., Naumann, G.O., 1988d. Optic disc, cup and neuroretinal rim size, configuration and correlations in normal eyes. *Invest. Ophthalmol. Vis. Sci.* 29, 1151–1158.
- Jonas, J.B., Papastathopoulos, K.I., 1996. Optic disc shape in glaucoma. *Graefes Arch. Clin. Exp. Ophthalmol.* 234 (Suppl. 1), S167–S173.
- Jonas, J.B., Kling, F., Gründler, A.E., 1997. Optic disc shape, corneal astigmatism, and amblyopia. *Ophthalmology* 104, 1934–1937.
- Klingenberg, C.P., 2008. MorphoJ. Faculty of Life Sciences, University of Manchester, UK. http://www.flywings.org.uk/MorphoJ_page.htm.
- Klingenberg, C.P., Monteiro, L.R., 2005. Distances and directions in multidimensional shape spaces: implications for morphometric applications. *Syst. Biol.* 54, 678–688.
- Kuhl, F.G.C., 1982. Elliptic Fourier features of a closed contour. *Comput. Graph. Image Process.* 18, 236–258.
- Manly, F.J., 1997. *Randomization, Bootstrap and Monte Carlo Methods in Biology*. Chapman & Hall, London.
- Morgan, J.E., Sheen, N.J., North, R.V., Choong, Y., Ansari, E., 2005a. Digital imaging of the optic nerve head: monoscopic and stereoscopic analysis. *Br. J. Ophthalmol.* 89, 879–884.
- Morgan, J.E., Sheen, N.J., North, R.V., Goyal, R., Morgan, S., Ansari, E., Wild, J.M., 2005b. Discrimination of glaucomatous optic neuropathy by digital stereoscopic analysis. *Ophthalmology* 112, 855–862.
- Neff, N.A., Marcus, L.F., 1980. *A Survey of Multivariate Methods for Systematics*. Privately Published, New York.
- Pablo, L.E., Ferreras, A., Fogagnolo, P., Figus, M., Pajarín, A.B., 2009. Optic nerve head changes in early glaucoma: a comparison between stereophotography and Heidelberg retina tomography. *Eye (London, England)*.
- Perez, S.I., Bernal, V., Gonzalez, P.N., 2006. Differences between sliding semi-landmark methods in geometric morphometrics, with an application to human craniofacial and dental variation. *J. Anat.* 208, 769–784.
- Rohlf, F.J., 2006. NTSYSpc, Version 2.21b. Exeter Software, Setauket, NY.
- Rohlf, F.J., 1990. Fitting curves to outlines. In: Rohlf, F.J., Bookstein, F.L. (Eds.), *Proceedings of the Michigan Morphometrics Workshop*. Special Publication Number 2. University of Michigan Museum of Zoology, Ann Arbor, pp. 167–177.
- Rohlf, F.J., 1998. On applications of geometric morphometrics to studies of ontogeny and phylogeny. *Syst. Biol.* 47, 147–158.
- Rohlf, F.J., 2008. tpsRelw version 1.46. Department of Ecology and Evolution, State University of New York at Stony Brook, Stony Brook. Available at: <http://life.bio.sunysb.edu/morph/>.
- Sanfilippo, P.G., Cardini, A., Hewitt, A.W., Crowston, J.G., Mackey, D.A., 2009. Optic disc morphology – rethinking shape. *Prog. Retin. Eye Res.* 28, 227–248.
- Sheen, N.J., Morgan, J.E., Poulsen, J.L., North, R.V., 2004. Digital stereoscopic analysis of the optic disc: evaluation of a teaching program. *Ophthalmology* 111, 1873–1879.
- Shrout, P.E., Fleiss, J.L., 1979. Intraclass correlations: uses in assessing rater reliability. *Psychol. Bull.* 86, 420–428.
- Slice, D., 1999. Morphheus et al. (beta version). Department of Ecology and Evolution, State University of New York, Stony Brook, New York.
- Slice, D., 2007. Geometric morphometrics. *Annu. Rev. Anthropol.* 36, 261–281.
- Stefánsson, M.O., Sigurdsson, T., Pampoulie, C., Daniëlsdóttir, A.K., Thorgilsson, B., Ragnarsdóttir, A., Gíslason, D., Coughlan, J., Cross, T.F., Bernatchez, L., 2009. Pleistocene genetic legacy suggests incipient species of *Sebastes mentella* in the Irminger Sea. *Heredity*.
- Zelditch, M.L., Swiderski, D.L., Sheets, D.H., Fink, W.L., 2004. *Geometric Morphometrics for Biologists*. Academic Press, London.