Anatomy and Pathology/Oncology

Pattern Recognition Analysis of Age-Related Retinal Ganglion Cell Signatures in the Human Eye

Nayuta Yoshioka,1,2 Barbara Zangerl,1,2 Lisa Nivison-Smith,1,2 Sieu K. Khuu,2 Bryan W. Jones,3 Rebecca L. Pfeiffer,3 Robert E. Marc,3 and Michael Kalloniatis1,2

1Centre for Eye Health, University of New South Wales (UNSW), Sydney, New South Wales, Australia
2School of Optometry and Vision Science, UNSW, Sydney, New South Wales, Australia
3Department of Ophthalmology, Moran Eye Center, University of Utah, Salt Lake City, Utah, United States

Correspondence: Michael Kalloniatis, Centre for Eye Health, Rupert Myers Building (South Wing), UNSW, Sydney 2052, Australia; M.Kalloniatis@unsw.edu.au.
Submitted: January 8, 2017
Accepted: May 5, 2017

PURPOSE. To characterize macular ganglion cell layer (GCL) changes with age and provide a framework to assess changes in ocular disease. This study used data clustering to analyze macular GCL patterns from optical coherence tomography (OCT) in a large cohort of subjects without ocular disease.

METHODS. Single eyes of 201 patients evaluated at the Centre for Eye Health (Sydney, Australia) were retrospectively enrolled (age range, 20–85); 8 × 8 grid locations obtained from Spectralis OCT macular scans were analyzed with unsupervised classification into statistically separable classes sharing common GCL thickness and change with age. The resulting classes and gridwise data were fitted with linear and segmented linear regression curves. Additionally, normalized data were analyzed to determine regression as a percentage. Accuracy of each model was examined through comparison of predicted 50-year-old equivalent macular GCL thickness for the entire cohort to a true 50-year-old reference cohort.

RESULTS. Pattern recognition clustered GCL thickness across the macula into five to eight spatially concentric classes. F-test demonstrated segmented linear regression to be the most appropriate model for macular GCL change. The pattern recognition-derived and normalized model revealed less difference between the predicted macular GCL thickness and the reference cohort (average ± SD 0.19 ± 0.92 and −0.30 ± 0.61 μm) than a gridwise model (average ± SD 0.62 ± 1.43 μm).

CONCLUSIONS. Pattern recognition successfully identified statistically separable macular areas that undergo a segmented linear reduction with age. This regression model better predicted macular GCL thickness. The various unique spatial patterns revealed by pattern recognition combined with core GCL thickness data provide a framework to analyze GCL loss in ocular disease.

Keywords: ganglion cells, image analysis (clinical), pattern recognition, optical coherence tomography, aging

The retinal ganglion cell (GC) is the final output cell of the retina, receiving and modulating input from bipolar cells and amacrine cells to code for complex visual information.1 While GCs are susceptible to a variety of disease processes,2–5 GC loss is also known to occur in the absence of identified disease as a part of aging.6–11 There is controversy regarding the spatial and temporal pattern of GC loss with age. Histologic studies have described GC loss to be linear,6,9,13 and a model for estimating the GC population derived from visual field sensitivity has likewise suggested a linear change.16 Studies using nerve fiber layer (NFL) thickness10,14,17–20 and ganglion cell layer (GCL) thickness1,2 from optical coherence tomography (OCT) as a measure of GC loss have likewise suggested linear loss with age. Gao and Hollyfield,7 however, found that while the GC loss appeared to be linear for the macular area, loss metrics in the peripheral retina appeared logarithmic. Closer inspection of histologic data from Harman et al.9 and Jonas et al.6 suggests that there is a large variation in GC count and limited losses until after middle age. Recent studies analyzing retinal21 and GCL thickness13 with the OCT in the macula likewise showed greater reduction in thickness after middle age. While it is now acknowledged that peripheral retina loses GCS faster than the macula,7,9 knowledge of the rate of macula change is limited. Inadequate understanding of age-related GC loss may confound the detection and diagnosis of age-related pathologies involving GC loss, namely glaucoma.15,18,22

OCT can provide high-density imaging and quantification of the retinal layers analogous to histologic studies,23–25 allowing large-scale in vivo profiling in a normal population. Thus, we sought to reinvestigate the pattern of normal age-related GC loss in the macula using OCT. Although numerous studies have quantified GC loss via NFL thickness10,14,17–20 and GCL thickness11,12 by OCT, few have analyzed these entities across temporal and spatial domains of the macula. Our study uses high-density macular cube OCT scanning to assess GC changes at 64 grid locations centered at the fovea, each grid 860 × 860 μm in size. To identify areas with similar age-related changes, we applied pattern recognition, a well-established technique for computationally clustering imaging data sets over N dimensions, in this case N =
7 age groups defined as decades. Pattern recognition visualizes complex data associations as memberships in statistically distinct theme classes and is traditionally used for satellite remote sensing analysis. 

This analysis can be applied in other fields and has previously been used to successfully cluster retinal cells into unique signature classes according to small molecule content. Pattern recognition has the distinct advantage of being able to analyze large data for multiobjective optimization and at the same time assess separability of the classes using statistical tests such as transformed divergence. 

We hypothesized that different areas within the macula could be allocated to distinct classes according to their temporal regression signature. While a gridwise analysis enables high spatial resolution, the discriminatory power may be reduced when analyzing data that are known to have high individual variability, such as GC density. 

Clustering analyses allow areas with statistically proven common features to be pooled together, which can improve the discriminatory power and allow more robust analysis of highly variable data. Furthermore, classwise data may facilitate the application of these data in future research by providing a framework with which macular areas can be classified and analyzed together. We further aimed to demonstrate the application and accuracy of these models for converting GCL thickness to an age equivalent, as well as a potential tool for in vivo estimation of GC count. This improves our understanding of normal human GC population dynamics and forms the basis for future investigations involving disease, particularly if areas of retinal space can be grouped into cohesive units.

Materials and Methods

Participants

Subjects were recruited retrospectively from patients who had been referred to the Centre for Eye Health for clinical imaging and assessments based upon risk factors or suspicion of ocular disease, but were subsequently found to have a healthy posterior pole in at least one eye. All patient records and imaging results were analyzed and reviewed by at least two experienced clinicians. Clinical data collected included visual acuity, retinal photography, Cirrus OCT scans (Macula 512 × 128 and Optic Nerve 200 × 200), intraocular pressure, and Spectralis OCT (Heidelberg Engineering, Heidelberg, Germany) macular scans for all participants spanned 30° × 25° and consisted of 61 B-scans spaced approximately 120 μm apart. Each B-scan image was constructed from averaging at least nine frames using the acquisition software’s Automatic Real Time (ART) mean function to reduce imaging noise. Scans with an overall quality score of less than 15 dB were excluded from the analysis. If multiple scans were available for the patient, the earliest scan meeting image quality criteria was selected. 

The measurement area was visualized on the HRA viewing module (version 6.3.4.0) consisting of an 8 × 8 grid (6880 × 6880 μm) centered on the fovea and the central horizontal line aligned over the line connecting the fovea and the middle of the optic nerve head (Figs. 1A, 1B). The GCL was automatically segmented throughout this area (Fig. 1C), and these were then manually checked and corrected for segmentation errors, defined as more than 50% difference in thickness compared to the surrounding area or any pronounced misclassification of the layers. Individual grids were excluded from the data if they contained segmentation errors that could not be corrected adequately (e.g., shadowing and distortion caused by blood vessels), or if 10% of the grid square fell outside the measurement area or was impinging on the optic nerve head.

Gridwise GCL Thickness Data Model

The cohort was separated into seven groups according to age by decade (Table 1). Normality of GCL thickness distribution for each decade group was checked with D’Agostino omnibus test (GraphPad Prism, version 6.0; GraphPad Software, La Jolla, CA, USA). The average GCL thickness was calculated for each grid location by decade (Supplementary Fig. S1). The GCL thickness data were converted to a log unit for two reasons: Linearization of the data and, as previous studies have shown, when GCL count is plotted against visual field sensitivity, another surrogate in vivo metric of GCL integrity, better correlation was achieved if GCL count was expressed as a log-unit.

Thus, the GCL thickness measurements were converted to decibel-micron units (dB, μm) using the following correlation:

\[
t_{\text{db}} = 10 \cdot \log_{10}(t_{\text{m}}),
\]

where \( t_{\text{db}} \) = GCL thickness in dB, \( t_{\text{m}} \) = GCL thickness in microns.

Linear regression and segmented linear regression analysis were conducted for the GCL thickness data for each grid location using GraphPad Prism (version 6.0) to obtain a gridwise rate of GCL thickness change as a function of age.

Pattern Recognition (Clustering) Model

A grayscale heat map (Fig. 1D, first image) was generated using gridwise GCL thicknesses in microns (Supplementary Fig. S1) to provide a range of pixel values between 0 and 255. Various scaling proportions were tested, with no difference found in the final classification results. Clustering of data requires appropriate “feature selection” to be conducted, as excessive numbers of classes do not aid classification if they are not separable and lead to unnecessary burden of the model. As a consequence, clustering is significantly affected by the applied strategy and statistical criteria. Therefore, we have adopted a clustering paradigm that has been well established in previous studies. Specifically, the data were analyzed with the tenets of the Declaration of Helsinki and ethics procedures put forward by the UNSW Australia HREA were observed for subject data collection.

Macular OCT

Spectralis OCT (Heidelberg Engineering, Heidelberg, Germany) macular scans for all participants spanned 30° × 25° and consisted of 61 B-scans spaced approximately 120 μm apart. Each B-scan image was constructed from averaging at least nine frames using the acquisition software’s Automatic Real Time (ART) mean function to reduce imaging noise. Scans with an overall quality score of less than 15 dB were excluded from the analysis. If multiple scans were available for the patient, the earliest scan meeting image quality criteria was selected. The measurement area was visualized on the HRA viewing module (version 6.3.4.0) consisting of an 8 × 8 grid (6880 × 6880 μm) centered on the fovea and the central horizontal line aligned over the line connecting the fovea and the middle of the optic nerve head (Figs. 1A, 1B). The GCL was automatically segmented throughout this area (Fig. 1C), and these were then manually checked and corrected for segmentation errors, defined as more than 50% difference in thickness compared to the surrounding area or any pronounced misclassification of the layers. Individual grids were excluded from the data if they contained segmentation errors that could not be corrected adequately (e.g., shadowing and distortion caused by blood vessels), or if 10% of the grid square fell outside the measurement area or was impinging on the optic nerve head.

The average GCL thickness was calculated for each grid location by decade (Supplementary Fig. S1). The GCL thickness data were converted to a log unit for two reasons: Linearization of the data and, as previous studies have shown, when GCL count is plotted against visual field sensitivity, another surrogate in vivo metric of GCL integrity, better correlation was achieved if GCL count was expressed as a log-unit. 

Thus, the GCL thickness measurements were converted to decibel-micron units (dB, μm) using the following correlation:

\[
t_{\text{db}} = 10 \cdot \log_{10}(t_{\text{m}}),
\]

where \( t_{\text{db}} \) = GCL thickness in dB, \( t_{\text{m}} \) = GCL thickness in microns.

Linear regression and segmented linear regression analysis were conducted for the GCL thickness data for each grid location using GraphPad Prism (version 6.0) to obtain a gridwise rate of GCL thickness change as a function of age.

Pattern Recognition (Clustering) Model

A grayscale heat map (Fig. 1D, first image) was generated using gridwise GCL thicknesses in microns (Supplementary Fig. S1) to provide a range of pixel values between 0 and 255. Various scaling proportions were tested, with no difference found in the final classification results. Clustering of data requires appropriate “feature selection” to be conducted, as excessive numbers of classes do not aid classification if they are not separable and lead to unnecessary burden of the model. As a consequence, clustering is significantly affected by the applied strategy and statistical criteria. Therefore, we have adopted a clustering paradigm that has been well established in previous studies. Specifically, the data were analyzed with the tenets of the Declaration of Helsinki and ethics procedures put forward by the UNSW Australia HREA were observed for subject data collection.
**Figure 1.** Illustration of ganglion cell analysis, conversion, and comparison. Ganglion cell layer (GCL) thickness measurements were obtained with the Spectralis OCT (A) identifying the macular and optic nerve head (ONH) center. (B) The default 64-square analysis grid of the Spectralis OCT was aligned to connect the two entities through the horizontal midline. Values for each grid location were based on average thickness values from (C) segmentation of the GCL identified between the purple and green lines. (D) Using the average GCL thickness for each decade of life, grid locations were converted into grayscale maps translating to distinct signatures. Grid locations with similar signatures were clustered into $n$ dimensional space into classes and a color code was assigned to each class to generate a theme map indicating grid locations that has statistically similar GCL thickness change over time.

**Conversion of GCL Thickness to an Age Equivalent**

The GCL thickness value of the entire cohort was converted to a 50 year old equivalent using the the regression model derived from linear and segmented linear models derived from grid-wise and pattern recognition (both raw and normalized data) analysis. This was compared to a 50 year old sub-cohort to determine the validity of each of the regression model.
unsupervised classification using ISODATA clustering for each decade subgroup (PCI Geomatica, Markham, ON, Canada) (Fig. 1D, second image) generating clusters of locations within the macula with similar change in GCL thickness with age (theme classes). ISODATA clustering is a specific form of K-means clustering (a migrating means methods) and aids with feature selection by automated splitting of high variance classes and merging classes with low separability. ISODATA clustering is not bound by a predefined number of classes, but allows the class numbers to be reduced or increased appropriately within a given range. The separability of identified theme classes was statistically verified using transformed divergence (DT). DT value ranges from 0 to 2, with 0 denoting complete separation. A value of >0.9 corresponds to a probability of correct classification of >98% and is commonly accepted as the cutoff for statistically significant separability for clustering studies. As well as recommended by the manufacturer of the software. Following classification and confirmation of separability, each distinct theme class was assigned a color for visualization in pseudocolor plot (Fig. 1D, third image). In the initial model, a single peripheral point was clustered together with the central four points corresponding to the fovea, which is inconsistent with an a priori assumption that the foveal pit is anatomically distinct from the rest of the retina. Thus, in developing the final model, the central four locations corresponding to the fovea were masked and the data were reanalyzed, resulting in the reassignment of the aforementioned peripheral point with no other change in the classification. The ISODATA algorithm and statistical analysis provided the maximum number of statistically separable classes. After identifying the highest number of statistically unique theme classes using ISODATA clustering (N = 8 theme classes), the K-means algorithm was used to restrict the number of classes stepwise down to the lowest separable number (N = 5) to further explore the effect of total number of theme classes on the provided model.

In a similar fashion to the gridwise analysis data, linear regression and segmented linear regression analysis were conducted on the eight theme classes to obtain a classwise rate of regression with age. F-test was conducted for each theme class to compare linear and segmented linear fit quality. The degree of freedom (df1) was defined as 1 for the numerator and 12 for the denominator (df2). The F-ratio was calculated for each theme class, defined as the quotient of the absolute sum of squares (SSR) of the linear regression divided by the absolute SSR of segmented linear regression. Cutoff for statistical significance was specified as >4.75 (equivalent P < 0.05). Paired t-tests were conducted for the absolute SSR of the theme classes to compare the fits of the two models throughout the entire measurement area.

Furthermore, to determine if the rate of GCL change is related to the baseline thickness, the GCL thickness data for each class were analyzed after being converted to proportional data by normalizing the data to the highest value. For the purpose of statistical assessment, the proportional data were further converted to a continuous variable by an ad hoc logit transform where \( t_{\logit} = \ln \left( \frac{tp}{1 - tp + \varepsilon} \right) \), where \( t' = \) the highest non-1 proportional GCL thickness value (dB, \( \mu m \)) from the data set, \( tp = \) GCL thickness (dB, \( \mu m \)) as a proportion, and \( t_{\logit} = \) proportional data after ad hoc logit transform.

**Table 1. Demographic Characteristics of the Study Population (n = 201)**

<table>
<thead>
<tr>
<th>Cohort</th>
<th>Average Age ± SD</th>
<th>Range</th>
<th>Rx ± SD, Diopters</th>
<th>IOP ± SD, mm Hg</th>
<th>Sex, Percent MF</th>
<th>Laterality, Percent OD-OS</th>
<th>Ethnicity, Percent Euro:Asian:Oth</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>All subjects</td>
<td>49.7 ± 15.8</td>
<td>20.2–84.9</td>
<td>−0.5 ± 1.8</td>
<td>15.9 ± 3.1</td>
<td>43.57</td>
<td>53.47</td>
<td>63:35:2</td>
<td>201</td>
</tr>
<tr>
<td>Cohorts by decade</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Second</td>
<td>25.4 ± 2.9</td>
<td>20–29</td>
<td>−1.3 ± 1.7</td>
<td>15.0 ± 2.8</td>
<td>37:63</td>
<td>63:37</td>
<td>36:64:0</td>
<td>27</td>
</tr>
<tr>
<td>Third</td>
<td>34.2 ± 3.0</td>
<td>30–39</td>
<td>−1.6 ± 1.7</td>
<td>14.6 ± 3.8</td>
<td>36:64</td>
<td>39:61</td>
<td>43:57:0</td>
<td>28</td>
</tr>
<tr>
<td>Fourth</td>
<td>44.9 ± 2.8</td>
<td>40–49</td>
<td>−0.8 ± 1.5</td>
<td>16.7 ± 2.6</td>
<td>34:66</td>
<td>39:61</td>
<td>59:36:5</td>
<td>44</td>
</tr>
<tr>
<td>Fifth</td>
<td>55.8 ± 2.8</td>
<td>50–59</td>
<td>−0.0 ± 1.0</td>
<td>16.6 ± 3.1</td>
<td>45:55</td>
<td>38:62</td>
<td>69:26:5</td>
<td>42</td>
</tr>
<tr>
<td>Sixth</td>
<td>65.5 ± 2.9</td>
<td>60–69</td>
<td>−0.2 ± 2.3</td>
<td>15.9 ± 3.1</td>
<td>51:49</td>
<td>51:49</td>
<td>74:26:0</td>
<td>35</td>
</tr>
<tr>
<td>Seventh</td>
<td>73.6 ± 2.8</td>
<td>70–79</td>
<td>+0.9 ± 1.3</td>
<td>15.8 ± 4.8</td>
<td>53:47</td>
<td>58:42</td>
<td>84:16:0</td>
<td>19</td>
</tr>
<tr>
<td>Eighth</td>
<td>82.0 ± 1.6</td>
<td>80–85</td>
<td>+0.8 ± 0.9</td>
<td>13.2 ± 2.6</td>
<td>83:17</td>
<td>83:17</td>
<td>83:17:0</td>
<td>6</td>
</tr>
<tr>
<td>50 year-old equivalent cohort</td>
<td>49.7 ± 2.9</td>
<td>46–55</td>
<td>−0.7 ± 1.4</td>
<td>16.9 ± 2.8</td>
<td>44:56</td>
<td>47:53</td>
<td>47:53:0</td>
<td>34</td>
</tr>
<tr>
<td>Curcio and Allen cohort*</td>
<td>35.5 ± 3.3</td>
<td>29–39</td>
<td>−1.7 ± 1.8</td>
<td>15.5 ± 2.7</td>
<td>32:68</td>
<td>42:58</td>
<td>39:58:5</td>
<td>31</td>
</tr>
<tr>
<td>Curcio and Allen’s cohort</td>
<td>33.8 ± 4.0</td>
<td>27–37</td>
<td></td>
<td></td>
<td>60:40</td>
<td>50:50</td>
<td></td>
<td>5</td>
</tr>
</tbody>
</table>

SD, standard deviation; Rx, habitual spectacle correction; OD, right eye; OS, left eye; Euro, European descent; Oth, ethnicity other than European or Asian.

* Cohort with an equivalent age distribution to Curcio and Allen’s cohort (P = 0.845, unpaired t-test).
Sum-of-squares $F$-test was conducted on the classwise normalized GCL thickness data to determine if all of the classes could be fitted to a single regression curve.

**Agewise Conversion: Validation of Developed Regression Models**

To validate the predictive value and accuracy of the regression models, GCL thicknesses for all subjects were converted to a 50-year-old equivalent using the previously developed regression models, specifically, the gridwise linear regression, classwise linear and segmented linear regression, and normalized segmented linear regression (Fig. 1E). Each resulting thickness plot was compared to the GCL thickness plots of a 50-year-old cohort ($n = 34$, Table 1) to assess for accuracy.

**GC Density per Volume**

GC density per GCL tissue volume (mm$^3$) was determined as previously described in Raza and Hood. In short, GC density per mm$^2$ for the horizontal and vertical meridians was linearly interpolated with polar coordinates for each of the 64 grid locations using currently available histologic data of GC density per area (mm$^2$). A group of age-equivalent subjects to the histologic data was selected from our cohort ($n = 31$, Table 1), and the gridwise GCL thickness was obtained from this group. GC density per mm$^2$ value was divided by the GCL thickness in mm to obtain the GC density in mm$^3$. Furthermore, gridwise estimate of the GC count for our age-equivalent subgroup was derived by multiplying the GC density per mm$^3$ to the average GCL thickness and grid area ($860 \times 860$ µm).

**RESULTS**

**Defining Cohort**

The majority of the study subjects were of European descent (61%) with roughly a third of the cohort Asian and with a slight bias toward females (57%; Table 1). The largest deviation from these averages was observed in the eighth decade, likely due to the small sample size (Table 1). GC density per mm$^2$ was interpolated from previous histologic data as described in Methods (Fig. 2A). From this, GC density per mm$^3$ (Fig. 2C) and the GC count (Fig. 2D) within the macula area were obtained, resulting in an estimated GC count for the total grid area of 4.68 $\times$ 10$^3$ cells, or 3.53 $\times$ 10$^3$ cells if restricted to the central 2.8-mm radius area.

**Gridwise Analysis of Age-Related Changes in the GCL Thickness**

GCL thickness was greatest at the paracentral macula with most areas exhibiting statistically significant rates of regression over time on linear regression analysis (Fig. 3A; mean ± SD: $-0.085 \pm 0.0311$ dB, µm) and reasonable coefficients of determination (mean $R^2$ ± SD: 0.648 ± 0.20). However, high variance was observed for both the rate of regression (range, $-0.003$ to $-0.141$ dB, µm per decade) and coefficient of determination (range $R^2$: <0.01-0.88) throughout different regions of the measurement area. In general, the central and paracentral points showed greater rate of regression with a bias toward the nasal quadrant and areas with the lowest statistical significance, and $R^2$ values were primarily located peripherally, although no clear trend could be identified (Fig. 3C). Segmented linear regression analysis for the gridwise data was unsuccessful as inflection point could be identified successfully for only 23 of the 64 grid locations, while the remaining 41 locations could not be adequately described using this model.

**Pattern Recognition Analysis**

Applying a cluster algorithm based on pattern recognition separated the investigated area into eight statistically distinct theme classes (implementing a >98% chance of correct classification: $D_T > 1.9$) representing macular locations with similar GCL thickness and change with age (Fig. 4A). These classes were arranged in concentric configurations, which were retained even when the number of theme classes was reduced (Fig. 5). The location provided theme maps (Figs. 4A, 5), which, when combined with the average thickness data (e.g., Supplementary Fig. S1), allowed the generation of theme class-derived data sets of GCL thickness change over time (e.g., Figs. 4B, 4C) for any theme-map combination. Theme classes plotted with the segmented linear regression exhibited negligible change in GCL thickness until the inflection point at middle to high age range (48.41–72.11 years of age; Table 2), after which greater regression rates were seen (average rate of regression before inflection: $-1.79 \times 10^{-3}$ dB, µm/decade; after inflection: $-3.45 \times 10^{-3}$ dB, µm/decade; Table 2). The average coefficient of determination for each of the regression models was found to be higher for segmented linear regression ($R^2 = 0.917$) compared to linear regression models ($R^2 = 0.663$; Table 2, Figs. 4B, 4C). This difference in coefficients of determination is to be expected, however, when comparing a simple model to a more complex one. To determine if the more complex model (segmented linear regression) is appropriate, $F$-test and paired $t$-test of the absolute SST was conducted. $F$-test exhibited $F$-ratios of greater than 4.75 (equivalent to $P < 0.05$) for the first four theme classes, demonstrating that segmented linear regression is a more appropriate model for these measurement areas (Table 3). The paired $t$-test did not reach statistical significance for all theme classes together ($P = 0.0543$), likely due to the high absolute SST exhibited by theme class 1, despite reaching significance by itself. When this class was excluded from the analysis, the relationship was found to be highly statistically significant ($P < 0.01$), thus suggesting that a segmented linear regression may be applied as a whole.

Sum-of-squares $F$-test was conducted on the normalized GCL thickness data to determine if all of the classes could be fitted to a single regression curve. Class 1 was analyzed independently on the a priori assumption that the foveal area, which it corresponds to, reflects an anatomically distinct entity. A model consisting of a different regression curve for each class was rejected by this analysis ($P = 0.983$), indicating that a single curve can adequately describe the observed regression pattern (Table 4).

**Age Equivalent Conversion**

To evaluate the predictive value of the aforementioned regression models, GCL thickness values of the entire cohort were converted to a 50-year-old equivalent using these models and compared to age-matched clinical data (Fig. 1E). Regression data for the linear curve were derived from both gridwise analysis (see Fig. 3) and pattern recognition analysis (Fig. 4B), while the regression data for segmented linear curve were derived from pattern recognition analysis (see Fig. 4C; Supplementary Table S1) and normalized data (Table 4), resulting in adjusted GCL thickness values (Supplementary Figs. S2B–E). The difference in GCL thickness of these converted values to the reference group (Supplementary Fig. S2A; Table 1) was found to be identical between the two linear regression models (average difference 0.67 $\pm$ 0.638 µm for both; Supplementary Figs. S2F, S2G). When segmented linear
regression was used, the disparity was reduced with both classwise and normalized analysis (0.50 \pm 0.446 and 0.42 \pm 0.344 \mu m, respectively; Supplementary Figs. S2H, S2I). Bland-Altman plots confirmed the above with the greatest bias introduced by the linear regression models (Fig. 6A; 0.51 \pm 1.11 \mu m, Fig. 6B; 0.51 \pm 1.10 \mu m: 95% limits of agreement) followed by normalized regression data (Fig. 6D; -0.30 \pm 0.61 \mu m) and then by classwise segmented linear model (Fig. 6C; 0.19 \pm 0.92 \mu m). While the normalized data showed larger amount of bias compared to the classwise segmented linear regression, the confidence interval was lower. When the disparity in GCL thickness between the classwise linear and...
FIGURE 3. Gridwise linear regression pattern (A). Pointwise regression data of ganglion cell layer thickness plotted over age for the 64 grid locations. (B) Magnified graph indicating for each grid location the rate of regression per decade in the top left corner and the coefficient of determination ($R^2$) in the top right and the relevant axis scales. (C) Heat map of the rate of regression for the 64 locations where red indicates highest rates of regression and green indicates lowest rates of regression. Error bars show standard deviation. Statistical significance of rate of regression (nonzero) $^*P < 0.05$ and $^{**}P < 0.01$, respectively.
segmented linear models was compared directly in a Bland-Altman plot as a function of eccentricity (Fig. 6E) and theme class (Fig. 6F), the segmented linear model specifically demonstrated superior accuracy for central retinal areas (i.e., classes 1–6), while at more peripheral grids, the two models were comparable (i.e., classes 7 and 8).

**DISCUSSION**

OCT is able to provide information analogous to histology in vivo, and the advancement in image analysis enables the GCL to be assessed independently from the inner plexiform layer (IPL), which arguably allows for better assessment of GC integrity as the GCL undergoes greater absolute thinning than the IPL. While a number of past studies have investigated age-related retinal tissue loss with OCT, to the best of our knowledge this is the first investigation that not only analyzed the gridwise GCL thickness changes but also clustered them into statistically separable theme classes according to temporal regression signatures. This study demonstrated that pattern recognition can be applied to retinal thickness measurement, with the advantage of allowing a large number of data points to be analyzed for multiobjective optimization over $n$ dimensions and, at the same time, statistically test the separability of the clusters.

Conventional gridwise analysis has the advantage of allowing analysis of an area without loss of spatial information. However, in a data set with high population variability, such as GCL thickness, the discriminatory power may be reduced. Our result demonstrated this to be the case; the spatial and temporal patterns of regression were not readily apparent with the gridwise regression data, and only a third of the measurement area could be modeled with segmented linear models was compared directly in a Bland-Altman plot as a function of eccentricity (Fig. 6E) and theme class (Fig. 6F), the segmented linear model specifically demonstrated superior accuracy for central retinal areas (i.e., classes 1–6), while at more peripheral grids, the two models were comparable (i.e., classes 7 and 8).
Ganglion Cell Layer (GCL) Thickness Regression Characteristic for the Eight Clusters Identified With Pattern Recognition Using the Linear Regression Model and Segmented Linear Regression Model. The First Segmented Linear Regression GCL Thickness Regression Rate is the Regression Rate Before Inflection, and the Second is for After Inflection ($R^2$, Coefficient of Determination)

### Linear Regression

<table>
<thead>
<tr>
<th>Theme Class</th>
<th>RGCL Thickness Regression, dB, μm/y</th>
<th>Intercept at Age 0</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$-1.04 \times 10^{-2}$</td>
<td>16.26</td>
<td>0.339</td>
</tr>
<tr>
<td>2</td>
<td>$-1.24 \times 10^{-2}$</td>
<td>17.78</td>
<td>0.701</td>
</tr>
<tr>
<td>3</td>
<td>$-9.61 \times 10^{-3}$</td>
<td>17.53</td>
<td>0.686</td>
</tr>
<tr>
<td>4</td>
<td>$-1.09 \times 10^{-2}$</td>
<td>16.78</td>
<td>0.753</td>
</tr>
<tr>
<td>5</td>
<td>$-1.02 \times 10^{-2}$</td>
<td>16.07</td>
<td>0.754</td>
</tr>
<tr>
<td>6</td>
<td>$-8.41 \times 10^{-3}$</td>
<td>15.42</td>
<td>0.733</td>
</tr>
<tr>
<td>7</td>
<td>$-7.64 \times 10^{-3}$</td>
<td>14.43</td>
<td>0.695</td>
</tr>
<tr>
<td>8</td>
<td>$-5.73 \times 10^{-3}$</td>
<td>13.64</td>
<td>0.645</td>
</tr>
<tr>
<td>Average</td>
<td>$-9.41 \times 10^{-3}$</td>
<td>15.99</td>
<td>0.665</td>
</tr>
</tbody>
</table>

### Segmented Linear Regression

<table>
<thead>
<tr>
<th>Theme Class</th>
<th>GCL Thickness Regression, dB, μm/y</th>
<th>Intercept at Age 0</th>
<th>Age of Inflection</th>
<th>GCL Thickness Regression, dB, μm/y</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>15.86</td>
<td>72.11</td>
<td>$-9.76 \times 10^{-2}$</td>
<td>0.933</td>
</tr>
<tr>
<td>2</td>
<td>$-2.37 \times 10^{-3}$</td>
<td>17.60</td>
<td>62.68</td>
<td>$-3.86 \times 10^{-2}$</td>
<td>0.995</td>
</tr>
<tr>
<td>3</td>
<td>$-1.83 \times 10^{-3}$</td>
<td>17.18</td>
<td>65.88</td>
<td>$-3.59 \times 10^{-2}$</td>
<td>0.987</td>
</tr>
<tr>
<td>4</td>
<td>$-3.79 \times 10^{-3}$</td>
<td>16.51</td>
<td>62.90</td>
<td>$-2.82 \times 10^{-2}$</td>
<td>0.954</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>15.71</td>
<td>50.86</td>
<td>$-1.84 \times 10^{-2}$</td>
<td>0.903</td>
</tr>
<tr>
<td>6</td>
<td>$-3.66 \times 10^{-3}$</td>
<td>15.23</td>
<td>63.00</td>
<td>$-2.07 \times 10^{-2}$</td>
<td>0.902</td>
</tr>
<tr>
<td>7</td>
<td>$-2.70 \times 10^{-3}$</td>
<td>14.24</td>
<td>63.00</td>
<td>$-2.04 \times 10^{-2}$</td>
<td>0.900</td>
</tr>
<tr>
<td>8</td>
<td>0</td>
<td>13.44</td>
<td>48.41</td>
<td>$-9.56 \times 10^{-3}$</td>
<td>0.765</td>
</tr>
<tr>
<td>Average</td>
<td>$-1.79 \times 10^{-3}$</td>
<td>15.72</td>
<td>61.11</td>
<td>$-3.45 \times 10^{-3}$</td>
<td>0.917</td>
</tr>
</tbody>
</table>

Table 3. The Absolute Sum Square and the F-Ratio for Comparison of the Two Regression Models, Linear and Segmented Linear Regression

<table>
<thead>
<tr>
<th>Theme Class</th>
<th>Absolute Sum Square</th>
<th>F-Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear</td>
<td>0.551</td>
<td>9.92*</td>
</tr>
<tr>
<td>Segmented</td>
<td>0.056</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.171</td>
<td>49.71*</td>
</tr>
<tr>
<td></td>
<td>0.005</td>
<td>22.59*</td>
</tr>
<tr>
<td></td>
<td>0.101</td>
<td>5.40*</td>
</tr>
<tr>
<td></td>
<td>0.087</td>
<td>2.54</td>
</tr>
<tr>
<td></td>
<td>0.007</td>
<td>2.72</td>
</tr>
<tr>
<td></td>
<td>0.006</td>
<td>3.05</td>
</tr>
<tr>
<td></td>
<td>0.047</td>
<td>1.52</td>
</tr>
</tbody>
</table>

* Statistically significant difference in linear and segmented linear model (F-Ratio of $>4.75$, corresponding to $P < 0.05$).

Analysis of our normalized GCL thickness found the rate of GCL change expressed as a percentage to be a constant over the whole macula area (except the fovea; Table 4). In other words, regression rate was proportional to the initial GCL thickness, which also follows a concentric pattern (Supplementary Fig. S1). Therefore, it may not be surprising that when macular areas are clustered according to their pattern of regression with age, they follow a similar concentric pattern. This is consistent with a previous study that found the NFL thickness rate of change to be proportional to the initial thickness. Analysis of our normalized GCL thickness found the rate of GCL change expressed as a percentage to be a constant over the whole macula area (except the fovea; Table 4). In other words, regression rate was proportional to the initial GCL thickness, which also follows a concentric pattern (Supplementary Fig. S1). Therefore, it may not be surprising that when macular areas are clustered according to their pattern of regression with age, they follow a similar concentric pattern. This is consistent with a previous study that found the NFL thickness rate of change to be proportional to the initial thickness. Further, to verify the pattern of cluster found, supplementary pattern recognition analysis was conducted on selected age groups (20–29, 40–49, and 60–69) of the cohort to determine if similar classes are still present through different age groups (Table 4).

<table>
<thead>
<tr>
<th>Retinal Area</th>
<th>GCL Thickness Regression (%)</th>
<th>Age of Inflection</th>
<th>GCL Thickness Regression (%)</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fovea (class 1)</td>
<td>0 72.11</td>
<td>$-0.609$</td>
<td>0.933</td>
<td></td>
</tr>
<tr>
<td>Nonfoveal classes combined</td>
<td>0 53.55</td>
<td>$-0.118$</td>
<td>0.867</td>
<td></td>
</tr>
</tbody>
</table>
Pattern Recognition Analysis of OCT Retinal Measurement

**FIGURE 6.** Comparison of linear and segmented linear model. Bland-Altman plot of difference ganglion cell layer (GCL) thickness between 50-year-old normative data and age-equivalent value converted with linear regression with gridwise analysis (A) and classwise analysis (B), classwise segmented linear (C) model, and normalized segmented linear regression (D) as a function of average GCL thickness. Note that class 1 areas, which constituted the foveal pits, are shown for reference but were excluded from analysis of the Bland-Altman plots as they are anatomically distinct from other retinal areas. The difference in linear (gridwise) and segmented linear (classwise) conversion values as a function of eccentricity from the fovea is shown in (E) and in relation to the theme classes shown in (F). Red fonts indicate areas where segmented linear model showed larger deviation from the 50-year-old equivalent value.

is expected due to the smaller sample size, it dropped below 1.9 on only one to three occasions for each subgroup, predominantly for the peripheral-most theme classes (data not shown). Furthermore, the concentric pattern persisted when clusters with similar characteristics were merged together with the restriction of the number of theme classes using the K-means algorithm (Fig. 5). Thus, these results lend support to the concentric pattern of theme classes found on
the initial ISODATA analysis as having biological meaning. Additionally, we present a schema with varying number of concentric theme classes, which allows macular area with statistically similar GCL characteristics to be analyzed together.

The concentric nature of GCL distribution and rate of change raises the question of the possibility of a uniform gridwise measurement area for quantifying the macular area. Arguably, measurement areas need to reflect the biology of the tissue being investigated, and the difficulties of applying a uniform grid are particularly evident in the foveal grids: It does not allow one to appreciate the large thickness change gradient present in this area (Figs. 1B, 1C). Previous histologic studies for photoreceptor32 and GCL37 count reduced the density of the sampling windows and increased window size with eccentricity, and a similar approach with OCT measurements needs to be explored.

The temporal regression pattern of macular GCL thickness was often modelled as a linear regression in previous studies,5,9,10,12,14,20 although there were some indications that they may follow a segmented linear regression.5,11 Comparison of the theme classes’ temporal regression pattern with F-test suggested that segmented linear regression was a more suitable model than linear regression for the determined classes. Alternatively, the GCL thickness regression can be expressed as a normalized percentage change, in which case, the one formula can be applied to the entire macular area (except the fovea). Both of these models demonstrated accurate conversion of GCL thickness value to a given age; the classwise model demonstrated greater accuracy (less bias) while the normalized model showed greater precision (smaller confidence interval). Depending upon the intended purpose, models with greater accuracy or precision may be desired: For instance, if converting a cohort’s GCL thickness data to a desired age to facilitate direct comparison to another cohort with a different age, the former may be preferable, although in cases such as repeated measurements on the same subject, the latter may be preferred. We therefore developed two alternate segmented linear regression models, each of which could be applied for different purposes. Combined with a spatial clustering schema with variable number of theme classes as shown in Figure 5, we present a framework with which future investigation of GCL can be conducted. In particular, it may be advantageous for applications in structure–function studies, as it may assist in the spatial translation of data to visual field, which has different spatial distribution of data. The GCL thickness measurement was represented as a decibel scale to further facilitate future comparative study, as visual field data are also expressed in decibel scale. While a linear scale has the advantage of direct applicability to clinical scenario and a uniform scale, it was shown by previous studies that a linear relationship exists between structure and function when decibel scale was used for both variables and better correlation may be achieved compared to linear scale.39

Gridwise GCL thickness measurement may be used as a tool for obtaining an estimate of the GC count in vivo for clinical and scientific purposes. While this has previously been investigated,40 a mismatch in age between the OCT and histology cohort potentially leads to overestimation of the GC count. To address this, we conducted a similar analysis with age-matched OCT and histology data57 (Table 1). The total GC count was estimated to be $3.55 \times 10^7$ cells within the central 2.8-mm radius, which was comparable to another OCT study ($3.81 \times 10^7$ cells)46 and histology ($3.69 \times 10^7$ cells)37-derived estimates. To further ascertain the impact of mismatch in age, the GC count was recalculated with a cohort with similar characteristics to that in Raza and Hood’s study (52.1 ± 9.14 years of age, $n = 135$, $P = 0.2024$, unpaired t-test), which did not considerably affect the estimate ($3.57 \times 10^7$ cells). This is perhaps not surprising given the nature of segmented linear regression, which showed slower regression until middle age. Disparities in GC count estimates may instead be attributed to other sources such as differences in OCT analysis methods, instrumentation, grid size, and analysis area. Additionally, the impact of nonneural elements (glia)37 and displaced amacrine cells must be considered.57,54–56 As acknowledged in the original paper,37 the histologic data utilized for this study also did not account for displaced amacrine cells and hence potentially overestimate GC density, although the impact within the macular region is expected to be low.57 Likewise, they did not account for the presence of glial cells within the GCL.53 Further investigation of displaced GC and glial cell density in the GCL of human retina may be required to refine the model further.

The temporal regression pattern for the GCL thickness was found to resemble the pattern of contrast sensitivity loss for high spatial frequency targets with age (Supplementary Fig. S4).58 While age-related loss of contrast sensitivity is commonly acknowledged, controversy exists regarding the extent to which optical and neural changes drive the losses.58–65 Similarity in the temporal pattern of GC loss in our data and contrast sensitivity loss suggests that age-related neural loss could be a fundamental source of visual sensitivity loss. On the other hand, comparison to visual field sensitivity with age showed a disparity in the rate of regression; visual field sensitivity within the central 10º regressed between -0.36 and -0.77 dB per decade,66 while the rates of GC loss were less when expressed as a linear model (approximately -0.1 dB, μm per decade, Table 2). These conflict with histology studies suggesting a direct relationship between GC count and visual field sensitivity,16 and a more recent investigation directly comparing GCL thickness measured with OCT likewise finding a direct relation between the two, albeit with a significant floor effect limiting the range where this is applicable.67 A possible explanation is a compensatory increase in the nonneuronal component, leading to a reduction of neural component density per unit volume.10,14,22 While there are studies that show the average GC soma diameter to decrease with glaucomatous neuropathy, which may lead to reduction in neural density,15,68–69 it does not alter significantly with age.9 Furthermore, recent studies have highlighted the limitation of conventional visual field strategy utilizing stimulus size not scaled for spatial summation area,70–72 a potential confounding factor for any such structure–function comparison. Thus, a structure–function study comparing GC density to appropriately scaled visual stimulus and further analysis of GC density per volume may be required to clarify the relationship.

Limitations

A limitation of the study includes the decreased sample size for the higher age group, notably the 8th decade group due to the higher prevalence of ocular disease in this age group. It was found, however, that excluding the 8th decade group from the analysis had negligible impact to the pattern recognition theme classes or the regression model (data not shown).

The variability of the GCL thickness measurement with the OCT is also a potential confounding factor. It was found by a previous study that the variability of the Spectralis OCT for GCL thickness is within 3 μm.73 While the variability does not alter with eccentricity, this is expected to have a greater impact on the peripheral areas than the central area due to the decreased initial thickness and is likely to impact the analysis of regression pattern peripheral locations, especially if rate of regression is proportional to the baseline thickness.

Finally, since this study utilized pattern recognition to cluster retinal areas, the representation of separable classes
reflects the specific criteria chosen to define separability. As such, application of different clustering algorithms may result in changes to the statistically supported distinction of individual groups or recognize additional subgroups. However, the current clustering paradigm had been established by previous studies, and we have further confirmed the separability of the clusters using a validated statistical method and cutoff value ($D_t > 1.9$). Using this paradigm, we identified a maximum of eight clearly distinct classes arranged in a concentric pattern. This was consistent with the GCL thickness pattern, which was also arranged concentrically and, given that GCL regression rate is proportional to the initial thickness, the concentric arrangement of the theme classes likely reflects the biology of the GCL rather than an artefact caused by the clustering process or criteria. The robustness of these classes could be further supported if they are reproducible using different measures of separability, such as $\Delta K$, which will be an important future step in this area once larger sample sizes are available.

CONCLUSIONS

Our study has demonstrated the advantage of pattern recognition for analyzing retinal OCT data, with the 64 grid locations rigorously grouped into a varying number of statistically separable theme classes and accentuating the temporal trend in decreasing GCL thickness. Specifically, we present two temporal regression models, one based on pattern recognition-derived theme classes and the other on normalized thickness data, and proposed a spatial clustering schema composed of a varying number of theme classes. These tools may form the basis of future investigations. For instance, either of the temporal regression models may be applied for conversion of GCL thickness data to a given age equivalent in a similar fashion to visual field regression data used in previous studies. The spatial theme class schema may be implemented in structure-function concordance study by allowing multiple measurement areas to be analyzed together and facilitate spatial translation of GCL data to visual field data points. A preliminary result on similar pattern recognition study on visual field revealed a similar spatial pattern to be present when stimulus size was adjusted for spatial summation area.

Acknowledgments

Supported by NHMRC 1033224; UNSW ECR Grant 2016 PS55430; NY was supported through a PhD scholarship provided by Guide Dogs NSW/ACT and Australian Postgraduate Award; Guide Dogs NSW/ACT is also a partner in the NHMRC grant and provided support for LNS; NIH NEI EY02576, EY015128, EY014800 Core Grant, an unrestricted grant from Research to Prevent Blindness to the Moran Eye Centre.

Disclosure: N. Yoshioka, None; B. Zangerl, None; L. Nivison-Smith, None; S.K. Khoo, None; B.W. Jones, None; R.L. Pfeiffer, None; R.E. Marc, None; M. Kalioniatis, None

References


