Evaluation of a Method for Estimating Retinal Ganglion Cell Counts Using Visual Fields and Optical Coherence Tomography

Ali S. Raza[1,2] and Donald C. Hood[1,3]

[1]Department of Psychology, Columbia University, New York, New York, United States
[2]Department of Neurobiology and Behavior, Columbia University, New York, New York, United States
[3]Department of Ophthalmology, Columbia University, New York, New York, United States

Correspondence: Donald C. Hood, Department of Psychology, 406 Schermerhorn Hall, 1190 Amsterdam Avenue, MC 5501, Columbia University, New York, NY 10027, USA; dch3@columbia.edu.

Submitted: October 27, 2014
Accepted: January 1, 2015

Citation: Raza AS, Hood DC. Evaluation of a method for estimating retinal ganglion cell counts using visual fields and optical coherence tomography. *Invest Ophthalmol Vis Sci*. 2015;56:2254–2268. DOI:10.1167/iovs.14-15952

Purpose. To evaluate the accuracy and generalizability of a published model that derives estimates of retinal ganglion cell (RGC) counts and relates structural and functional changes due to glaucoma.

Methods. Both the Harwerth et al. nonlinear model (H-NLM) and the Hood and Kardon linear model (HK-LM) were applied to an independent dataset of frequency-domain optical coherence tomography and visual fields, consisting of 48 eyes of 48 healthy controls, 100 eyes of 77 glaucoma patients and suspects, and 18 eyes of 14 nonarteritic anterior ischemic optic neuropathy (ION) patients with severe vision loss. Using the coefficient of determination R^2, the models were compared while keeping constant the topographic maps, specifically a map by Garway-Heath et al. and a separate map by Harwerth et al., which relate sensitivity test stimulus locations to corresponding regions around the optic disc. Additionally, simulations were used to evaluate the assumptions of the H-NLM.

Results. Although the predictions of the HK-LM with the anatomically-derived Garway-Heath et al. map were reasonably good ($R^2 = 0.31$–0.64), the predictions of the H-NLM were poor ($R^2 < 0$) regardless of the map used. Furthermore, simulations of the H-NLM yielded results that differed substantially from RGC estimates based on histology from human subjects. Finally, the value-added of factors increasing the relative complexity of the H-NLM, such as assumptions regarding age- and stage-dependent corrections to structural measures, was unclear.

Conclusions. Several of the assumptions underlying the H-NLM should be revisited. Studies and models relying on the RGC estimates of the H-NLM should be interpreted with caution.

Keywords: glaucoma, retinal ganglion cells, RGC, optical coherence tomography, OCT, visual fields, VF, standard automated perimetry, SAP, structure, function

The structure–function relationship between retinal ganglion cell (RGC) density (structure) and light sensitivity (function; i.e., behavioral thresholds) is not fully understood, nor is the change in this relationship when RGCs undergo atrophy as the result of neurodegenerative diseases such as glaucoma. Early attempts to assess this relationship in human subjects[1–3] analyzed postmortem histology to measure structure, whereas more recent work has used optical coherence tomography (OCT), an in vivo imaging technique[4] that allows for direct measurements of the anatomy of the human eye. These imaging data also can be used to identify early structural changes in the human retina that are consistent with glaucomatous neurodegeneration.[5,6] In particular, this technique allows for visualization of the individual layers of retinal anatomy in a manner similar to histology,[7,8] allowing the opportunity to test structure–function models by using OCT data from the eyes of a large number of human individuals rather than a smaller number of postmortem human donor eyes.

To date, there are three predictive structure–function models pertaining to glaucoma that have been applied to OCT data: the models of Harwerth et al.,[9,10] Hood and Kardon,[11] and Wollstein et al.[12] (Malik et al.[13] provides an overview of several important structure–function models, including those using structural measures other than OCT.) The relatively new Wollstein et al.[12] model focuses primarily on the clinical relevance of the relationship between visual sensitivity and OCT; this model is geared toward determining the association between statistically significant visual field (VF) loss and statistically significant structural loss. On the other hand, the Harwerth et al.[9,10] model and the Hood and Kardon[11] model both attempt to describe the nature of the structure–function relationship in glaucoma.

Harwerth et al.[14] originally developed a model based on a mechanically-induced experimental model of glaucoma in rhesus macaques. These monkeys were trained to perform VFs, and these behavioral data were compared with RGC density obtained from analysis of postmortem histology. Along with previous histological studies in human subjects,[1–3] the work of Harwerth et al.[10] was a fundamental step in advancing the discourse regarding glaucomatous structure–function relationships. In particular, Harwerth et al.[10] proposed a

quantitative model, stating that when both measures were plotted on a logarithmic scale, visual sensitivity was linearly related to ganglion cell density with a nonzero intercept that varied with retinal eccentricity, or rather that the relationship was nonlinear when both structure and function were expressed in linear units. (Hereafter, the term "linear" is used only to describe a relationship that is linear when both measures are expressed in linear units.) The Harwerth et al.[10] nonlinear model (H-NLM) was later modified to describe OCT data from human subjects by using the thickness of the retinal nerve fiber layer (RNFL) around the optic disc as a proxy for the number of RGC axons. Later iterations of the model included a correction for age[15,16] and stage of disease.[17] One attractive aspect of the H-NLM is that it converts both visual sensitivity and OCT-derived RNFL thickness to estimates of numbers of RGC bodies and axons, particularly appealing as glaucoma is a disease of the RGCs.

Another approach was that of Hood, Kardon, and colleagues,[11,18–22] who directly compared visual sensitivity to OCT-derived RNFL thickness and proposed a linear relationship with a nonzero intercept (residual nonneuronal RNFL thickness) when both measures are expressed in linear units. Thus, a 50% loss in visual sensitivity compared with age-matched normative data should be associated with a 50% loss in the neuronal component of the RNFL. However, unlike the H-NLM, the Hood and Kardon linear model (HK-LM) does not predict that RNFL thickness will have a strong association with greater-than-normal (>100%) visual sensitivity and does not correct structural measurements for age or stage of disease. Thus, the HK-LM has considerably fewer parameters than the H-NLM, although it is also less ambitious in its scope, as it does not attempt to relate either functional or structural measures to estimates of RGCs.

Newer frequency-domain OCT (fdOCT)[23] has a faster acquisition rate, as well as higher axial resolution. Thus, fdOCT data should theoretically provide a more accurate measurement of RNFL thickness than older time-domain OCT (tdOCT). Additionally, although circle scans around the optic disc acquired from tdOCT are usually not perfectly centered, potentially affecting the accuracy and precision of RNFL thickness measurements,[11,24,25] the volume scans produced by the fdOCT allow for postacquisition adjustment of the location of a derived circle scan. Several independent groups have compared the predictions of the HK-LM to the RNFL thickness measurements derived from fdOCT data, and these studies[26–29] suggest that the HK-LM generalizes fairly well to fdOCT data. Although the H-NLM has been applied to fdOCT data, these studies[30–34] used the RGC estimates derived from the H-NLM for further modeling, rather than formally assessing the validity or accuracy of the model itself.

Because parameters have been added to the H-NLM over time, it is particularly important to assess the generalizability of the model. Harwerth et al.[10] previously applied the H-NLM to two different independent datasets and concluded that the H-NLM could "be generalized to other patient populations with equal results." Moreover, Harwerth et al.[10] directly compared the performance of the H-NLM to the HK-LM and concluded that although the performance of the models was "very similar, both in terms of the accuracy and the precision," the HK-LM was "less precise." However, given the importance of the H-NLM, there are several aspects of this validation that warrant further investigation.

First, the generalizability of a model can be determined only if the parameters are fixed before application to an independent dataset. Whereas the Harwerth et al.[10] validation of the H-NLM did not involve the addition of any new explicit parameters, there is at least one important implicit parameter that was changed before the H-NLM was applied to new datasets: a new topographic "map" relating measurement of axons near the optic disc to RGC bodies throughout the retina. Additionally, the previous validation study of the H-NLM compared the H-NLM to the HK-LM using different topographic maps for the two models. Finally, the validation study used older tdOCT data instead of the fdOCT data used in subsequent studies[30–34] by other groups.

The impact of the work by Harwerth et al.[10] cannot be underestimated. In the past few years, many publications[30,32–37] from outside groups have used the H-NLM as a basis for further modeling. In fact, regarding a new model that uses the H-NLM to estimate RGC counts, a recent review[38] has said that "[i]ts use in clinical trials may potentially overcome the limitations of currently available conventional parameters." Given this rise in prominence of the H-NLM, the purpose here was 3-fold: to apply the H-NLM to an independent fdOCT dataset and assess its generalizability, particularly with regard to the impact of different topographic maps; to compare the performance of the H-NLM and HK-LM using appropriate statistics, while controlling for differences in the topographic maps; and to use simulations to explore the underlying assumptions of the H-NLM with a view toward assessing their validity and accuracy.

Methods

Participants

Data were collected from three groups: 48 eyes of 48 controls (age = 51.4 ± 7.4 years [mean ± SD]) with healthy vision, 100 eyes of 77 glaucoma patients and suspects (age = 57.2 ± 12.0 years [mean ± SD]) with mostly mild to moderate vision loss, and 18 eyes of 14 nonarteritic anterior ischemic optic neuropathy (ION) patients (age = 62.9 ± 10.8 years [mean ± SD]) with severe vision loss.

Healthy control eyes were included based on the following criteria: spherical refraction between −6.0 and +3.0 diopters, IOP 21 mm Hg or lower, axial length between 22 and 26 mm, a normal clinical examination, and normal VFs. Individuals were excluded if they had a history of ocular disease or a family history of glaucoma. Controls were enrolled prospectively as part of a previous study.[39] The glaucomatous group consisted of patients in whom at least one eye exhibited glaucomatous optic neuropathy, defined based on stereophotography evaluation by glaucoma specialists using the following criteria: focal or diffuse neuroretinal rim thinning, focal or diffuse RNFL loss, or an intereye vertical cup-to-disc ratio asymmetry greater than 0.2 not explained by differences in disc size. All eyes had open angles as viewed during gonioscopic examination. Finally, the ION group was defined based on an assessment of history, clinical examination, and VFs by an experienced neuro-ophthalmologist. The ION eyes were included only if at least 6 months had elapsed after an acute event. For both the glaucoma and ION groups, consecutive patients were enrolled retrospectively based on the availability of test data. Patients with cataracts, a history of ocular surgery, or a history of any other ocular or neurological diseases that could affect structural or functional measures were excluded. For all three groups, structural and functional measures were required to be within 1 year of one another (time between tests = 56 ± 7 days [mean ± SD]).

Written, informed consent was obtained from all participants. Procedures followed the tenets of the Declaration of Helsinki, and the protocol was approved by the institutional review boards of Columbia University and the New York Eye and Ear Infirmary.

Investigative Ophthalmology & Visual Science

Function (Light Sensitivity): VF

All individuals were tested with standard automated perimetry (24-2 SITA Standard[40] protocol, Humphrey 750i Visual Field Analyzer; Carl Zeiss Meditec, Inc., Dublin, CA, USA). Subjects were required to have fixation losses of 33% or less, false-negatives of 33% or less, and false-positives of 15% or less. The false-negative requirement was not used for eyes with a mean deviation (MD) worse than −10 dB, because severe loss is associated with a higher reported rate of false-negatives. The MD (mean ± SD) for the control group was −0.2 ± 0.9 dB, the glaucomatous group −4.4 ± 5.6 dB, and the ION group −18.6 ± 7.3 dB.

Structure: fdOCT

All individuals were also tested using fdOCT (3D-OCT 1000/2000; Topcon Medical Systems, Inc., Oakland, NJ, USA) with the volume (cube) scan protocol (6 × 6 mm, 128 horizontal B-scans with 512 A-scans each) with an internal fixation target positioned to center the optic disc in the image frame. Scans with poor fixation and blink artifacts were rejected. The thickness of retinal layers was determined using a previously-validated segmentation algorithm,[41] which was manually corrected as necessary[42,43] based on the performance of the automated algorithm. The manual correction was done by individuals masked to the classification of each eye (i.e., healthy or glaucomatous). In particular, the thickness of the RNFL was determined from a circle (3.45-mm diameter) extracted from the volume scan. The circle was manually centered post acquisition based on the location of the end of Bruch's membrane using simultaneous views of an en face summed-intensity projection image ("shadowgram"), as well as perpendicular slices (extracted "B-scans").

Topographic Maps Relating Structure and Function

The H-NLM derives separate estimates of RGC counts based on both the VF and OCT data. These estimates must be related by a topographic map, which specifies the spatial correspondence between a set of VF test stimulus locations (and the resulting RGC body estimate) and a particular region around the optic disc (and the resulting RGC-axon estimate). Figure 1 shows six topographic maps that have been used to relate structure and function. The H-NLM originally used the Harwerth et al.[44] 2007 map (Fig. 1C). In their validation study, Harwerth et al.[10] revised this map to the Harwerth et al.[10] 2010 map (Fig. 1D) and in a subsequent study to the Wheat et al.[17] 2012 map (Fig. 1E). Both of these newer maps relate the superior and inferior hemifields within approximately ±27° of visual angle from the fovea to a very wide region of the disc, including large portions of the nasal half of the disc. However, the nasal half of the disc has a large degree of axonal input from the peripheral retina outside the central approximately ±27° of visual angle (Fig. 1G, middle), which is not sampled by the 24-2 VF test stimulus locations.

Unlike the H-NLM, the HK-LM used only the Garway-Heath et al.[45] 2000 map (Fig. 1A), developed by tracing local defects and therefore based on anatomy. The Garway-Heath et al.[46] 2002 map (Fig. 1B) was a subsequent revision by the same group, but it was geared toward a structural measure other than the OCT. Both of these maps have shown generally good agreement with computational models describing the paths of the RGC axons in the human retina.[47–50]

Simulations of H-NLM

Because of the complexity of the H-NLM, it is difficult to disassociate the impact of each parameter and to examine the individual assumptions. Therefore, in addition to the data previously described, normative data for a range of ages were simulated for both the VF and OCT to better explore the H-NLM.

Average sensitivity values for each test stimulus location of the VF were generated for controls of varying ages using the normative database of the perimeter. For each test stimulus location of the 10-2, 24-2, and 30-2 VF test patterns, the average normal sensitivity in decibels can be derived from a set of linear equations, as follows:

$$\overline{s_c} = m_{sai}a + b_{sai}, \quad (1)$$

where the averaged sensitivity $\overline{S_C}$ in decibels for a group of age-matched normative individuals is given by a linear equation; a is age in years; m_{sai} is the location-dependent slope of declining sensitivity with age in decibels per year; and b_{sai} is the location-dependent intercept in decibels.

The exact values for m_{sai} and b_{sai} are proprietary (Carl Zeiss Meditec, Inc.), but are similar to previously published values (n = 140 eyes, age range, 20–80 years, 1 testing center).[51] The manufacturer values were used because of the greater sample size and age range (n = 422 eyes, age range, 17–89 years, 10 testing centers),[52] and in particular because values for locations closer to the fovea also were available.

Average RNFL thickness measurements for the entire optic disc also were generated for controls of varying ages using the equation published by Harwerth et al.[15] as follows:

$$\overline{t_c} = -0.3a + 110.0, \quad (2)$$

where $\overline{t_c}$ is the average control RNFL thickness in microns averaged around the entire optic disc, and a is age in years.

The average loss of 0.3 μm per year of RNFL thickness (n = 55 eyes) in the study by Harwerth et al.[15] is similar to a larger study[53] that reported 0.2 μm per year (n = 328 eyes).

Data Analysis and Evaluation

Data were analyzed using custom code written in MATLAB (version 2014a; MathWorks, Inc., Natick, MA, USA). Details of the H-NLM and HK-LM, including relevant equations, can be found in the Supplementary Material. To evaluate the models, the coefficient of determination R^2 was used, designated as R^2_{cd}. Occasionally, squared Pearson correlation coefficients were calculated, designated as R^2_P. Note that, unlike R^2_P, the coefficient of determination R^2_{cd} does not require a linear model and also can be negative based on the following equation:

$$R^2_{cd} = 1 - \frac{\sum_i (y_i - \hat{y}_i)^2}{\sum_i (y_i - \bar{y})^2}, \quad (3)$$

where the coefficient of determination R^2_{cd} is defined as 1 minus the ratio between the residual-sum-of-squares (observed value y_i for each data point i minus the corresponding model prediction $\hat{y}_i$, quantity squared) and the total-sum-of-squares (observed value for each data point y_i minus the mean of all observed values $\bar{y}$, quantity squared).

A value of R^2_{cd} = 1 indicates that the model accounts for the variability of the data perfectly, R^2_{cd} = 0 indicates that the model performs the same as a "null" model using merely the mean of the data plotted on the y-axis (e.g., the RNFL thickness), and $R^2_{cd} < 0$, indicates that the model performs

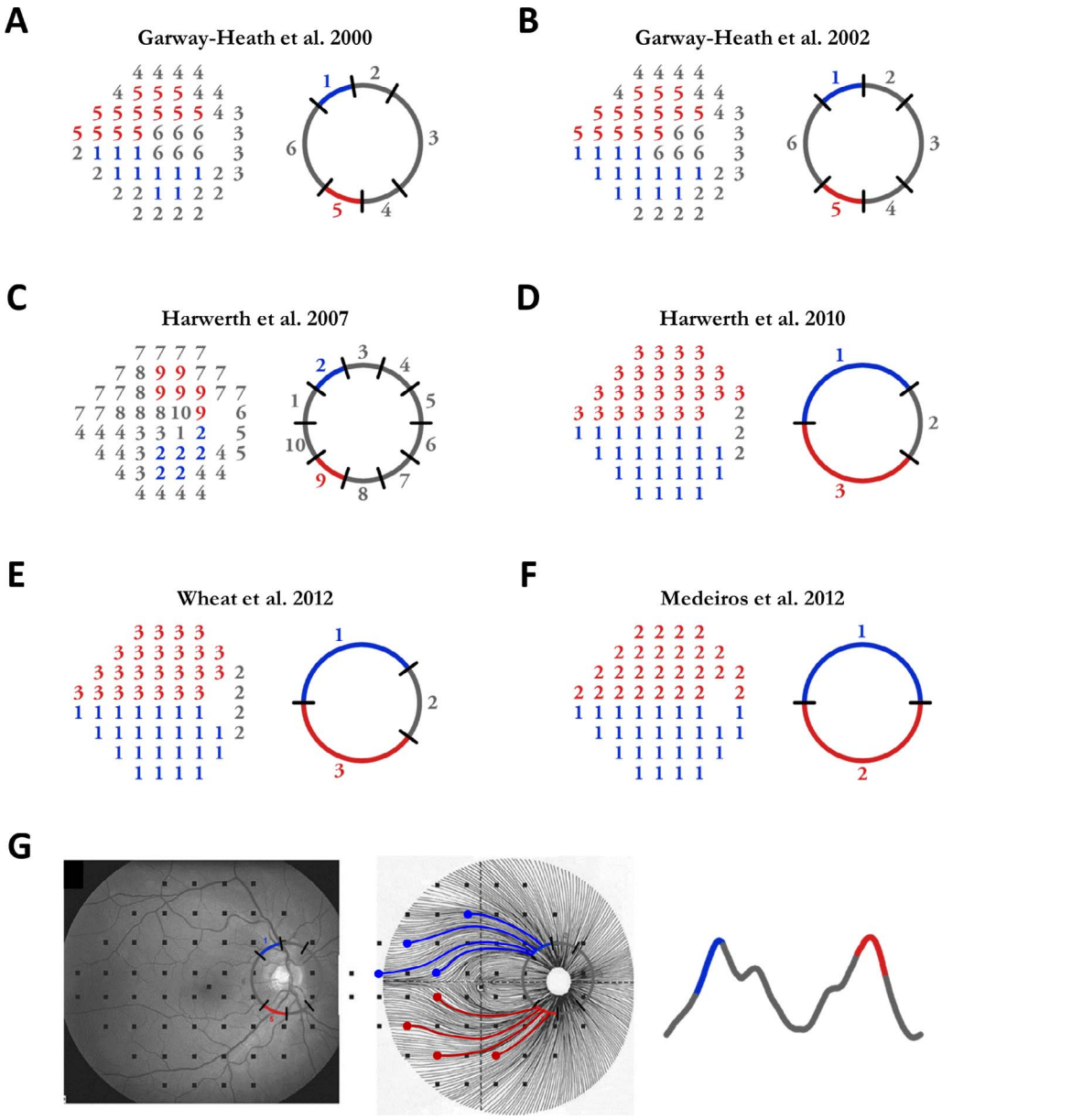

Figure 1. (**A**–**F**) Several topographic maps relating 24-2 VF stimulus locations to regions around the optic disc. (**G**) The Garway-Heath et al.[45] 2000 topographic map illustrated by (**A**) the test stimulus locations of the 24-2 VF superimposed on a photo of the fundus (*left*), the path of hypothetical RGC axons from test stimulus locations to the optic disc based on a schematic of RGC axons (*middle*), and the RNFL thickness profile around the optic disc averaged for all healthy control eyes (*right*).

worse than a null model. Note that when the same data are used to both generate and test a linear regression line, then the R^2_P will be equal to the R^2_{cd}; however, one can use the R^2_{cd} to assess a model derived from an independent dataset, which may be a more intuitive explanation for why the R^2_{cd} can be less than zero.

In the earlier work of Harwerth et al.,[10] the root-mean-squared-error (RMSE) was also used to evaluate the models and is defined as follows:

$$\varepsilon_{RMSE} = \sqrt{\frac{1}{n}\sum_i (y_i - \hat{y}_i)^2}, \quad (4)$$

where ε_{RMSE} is the RMSE; n is the total number of data points; y_i is the observed value for each data point i; and $\hat{y}_i$ is the corresponding model prediction.

Results

Test of the Models

To compare the performance of the H-NLM and HK-LM, the latest sets of published parameters were used for both models except as otherwise indicated. Figure 2A shows the results of the H-NLM in the same form as it is commonly plotted (e.g., Figs. 10A, 11A in the review by Harwerth et al.[10]). In general, there is agreement between the RGC predictions from both function (*x*-axis) and structure (*y*-axis) for the controls (green squares), glaucoma patients and suspects (black circles), and ION patients (red squares). Although there is a tendency for the points to fall below unity (gray line), the points are well correlated ($R^2_P = 0.90$). In the past, similar results have been used to argue in favor of the validity of the H-NLM. However, when the total RGC counts for an entire eye were determined

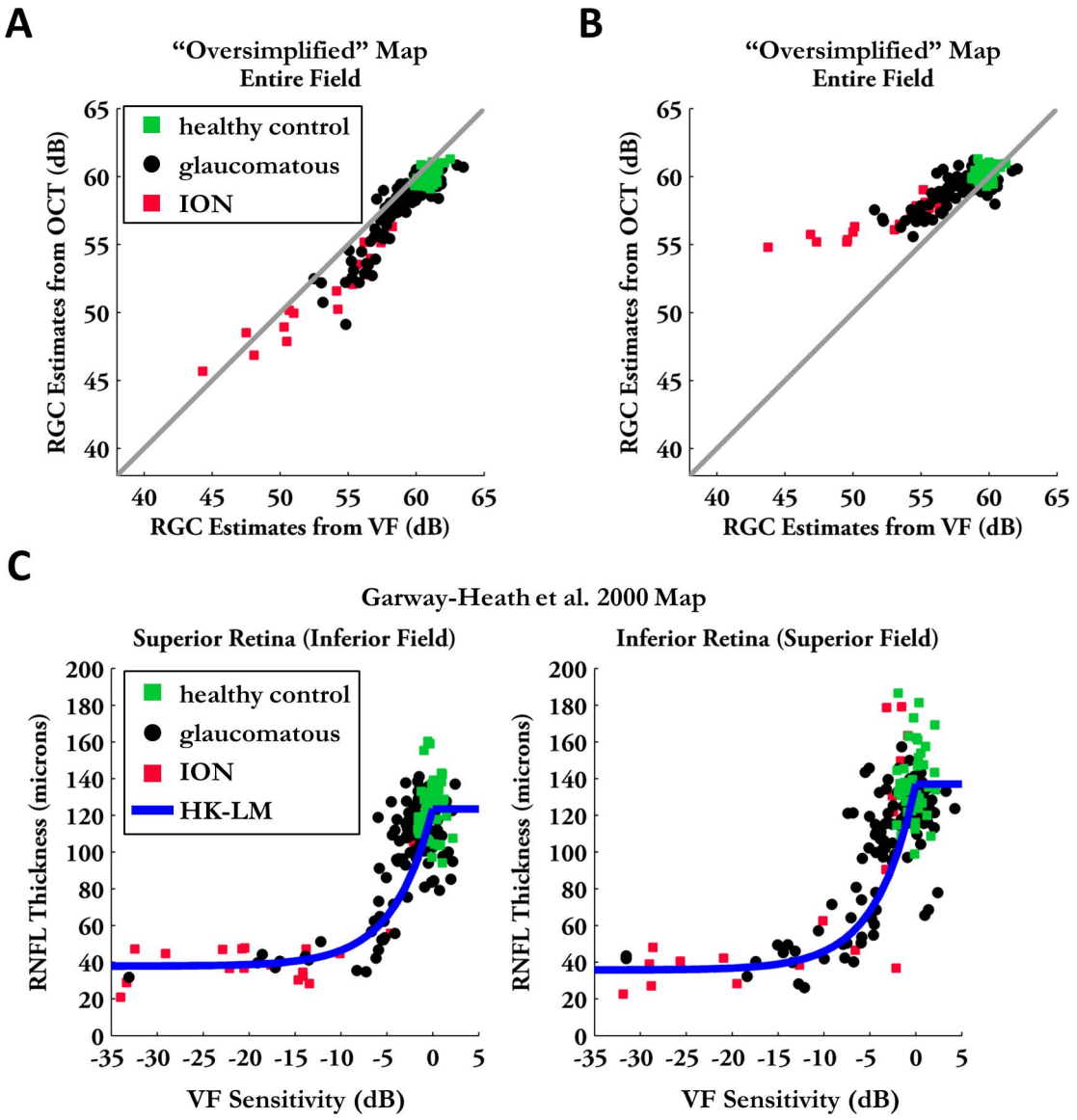

Figure 2. (**A**) Estimated RGC counts determined from VF (*x*-axis) and OCT (*y*-axis) data based on the H-NLM for controls (*green squares*), glaucoma patients and suspects (*black circles*), and ION patients (*red squares*). (**B**) The same data as in (**A**) when using the H-NLM with an older set of parameters and without the disease-stage correction. (**C**) The same data as in (**A**) plotted for the superior retina (*left*) and inferior retina (*right*) using the Garway-Heath et al.[45] 2000 map with the HK-LM (*blue*) superimposed.

by either Harwerth et al. (e.g., Ref. 15) or in further work by other groups, it is generally assumed that the entire region around the optic disc is related only to the region within the 24-2 VF test pattern (approximately ±27° of visual angle from the fovea), although, of course, the disc receives input from the entire retina. We refer to a map that relates the entire disc to the region covered by the 24-2 VF test pattern as an "oversimplified" map. Significantly, even though the agreement (as in Fig. 2A) between function-derived and structure-derived RGC estimates has been offered as evidence for the validity of the H-NLM, the RGC estimates, if accurate, should not agree when using the oversimplified map. This predicted disagreement can be seen when using older parameters (see Discussion) for the H-NLM, as shown in Figure 2B, where the RGC estimates based on the OCT are consistently greater than those based on the VF.

Using the Garway-Heath et al.[45] 2000 map, Figure 2C shows the prediction of the HK-LM (blue line) for both the inferior (left, $R^2_{cd} = 0.59$) and superior (right, $R^2_{cd} = 0.46$) hemifields for the same data shown in Figures 2A, 2B (although note the change in axes and units in Fig. 2C). The values for the parameters t_c and t_b (Equation S33; see the Supplementary Material for Equations S1–S34) for the HK-LM were set based on the controls and ION patients. If previously published[22] values are used for these parameters, the fit of the HK-LM is similar for both the inferior ($R^2_{cd} = 0.64$) and superior ($R^2_{cd} = 0.52$) hemifields; however, prior values do not exist for other topographic maps. Therefore, for the remaining analyses (unless otherwise indicated), both the H-NLM and HK-LM are assessed using only the glaucoma patients and suspects (black circles) and the subsequent R^2_{cd} values are based only on these data.

A direct comparison of the two models should use the same topographic map, so the models were tested using the Wheat et al.[17] 2012 map, which is very similar to the Harwerth et al.[10] 2010 map (see Methods and Fig. 1). The results can be seen in

Investigative Ophthalmology & Visual Science

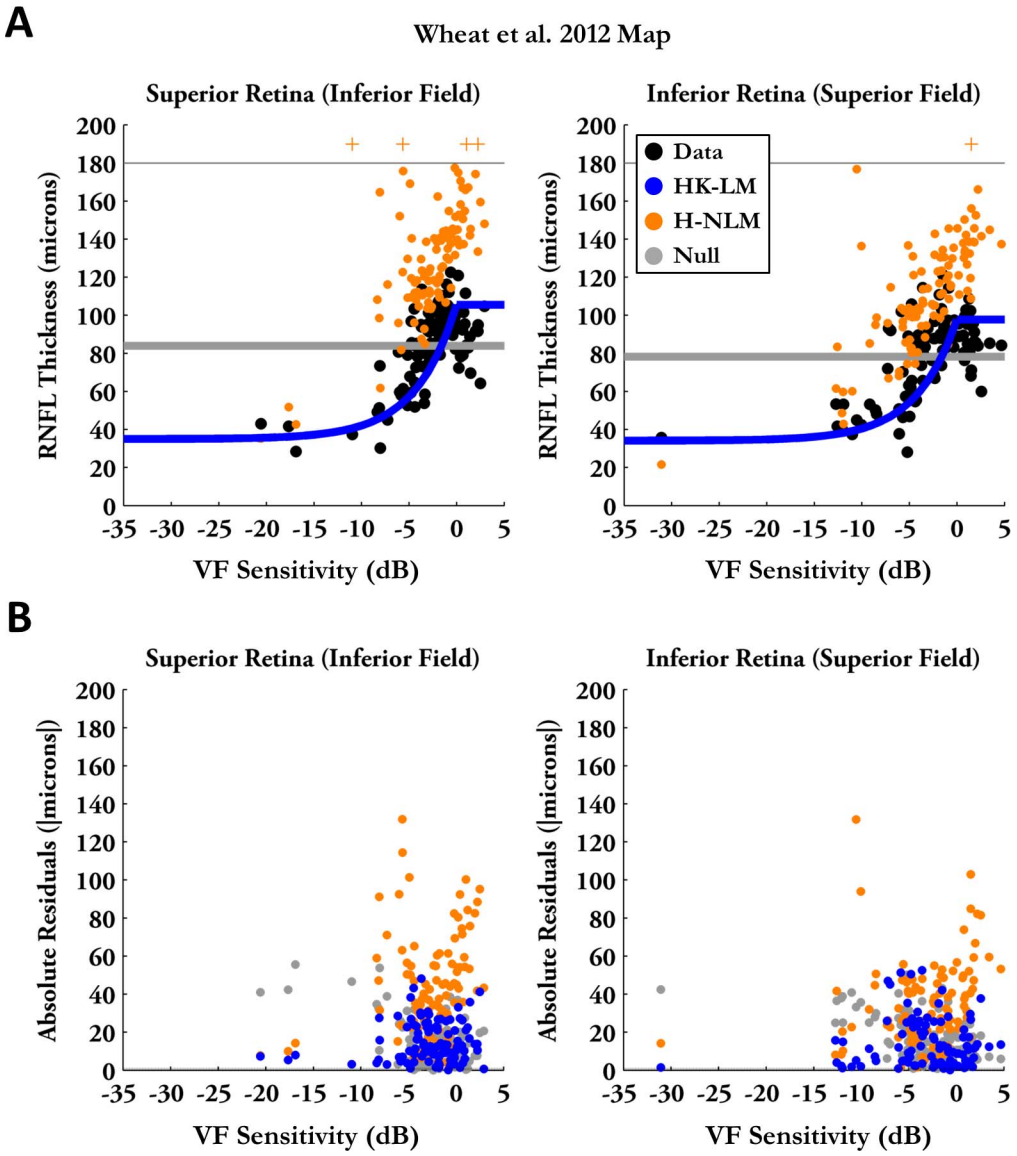

FIGURE 3. Structure and function related using the Wheat et al.[17] 2012 topographic map. The total deviation from age-matched normative data for the VF sensitivity is on the *x*-axis. (**A**) The RNFL thickness for the region corresponding to the VF, based on a topographic map, is on the *y*-axis. The predictions of both the HK-LM (*blue*) and the H-NLM (*orange*) for the data from the glaucoma patients and suspects (*black circles*) along with a "null" model (*gray horizontal line*). Outliers (*plus symbols*) are pinned to the top of the graph. (**B**) The absolute residuals of each of the models from the data (*black circles*) in (**A**).

Figure 3A, where the data from the glaucoma patients and suspects are shown as the black circles. The predictions of the HK-LM are shown as the solid blue line and the predictions of the H-NLM by the orange circles. The H-NLM predictions cannot be plotted as a simple function because the H-NLM includes individualized corrections for age and disease stage. (Note that outliers above the range of the *y*-axis are pinned to the top of the graph as + symbols.) The HK-LM (blue line, $R^2_{cd} = 0.15$ inferior hemifield, $R^2_{cd} = 0.14$ superior hemifield) describes the data from the glaucoma patients and suspects (black circles) better than the predictions of the H-NLM (orange circles, $R^2_{cd} = -6.86$ inferior hemifield, $R^2_{cd} = -3.18$ superior hemifield).

The absolute residuals of both models can be compared qualitatively in Figure 3B, where, as reflected in the R^2_{cd} values, the H-NLM residuals (orange circles) have a tendency to be larger than the HK-LM residuals (blue circles). When both models are plotted against the data using the anatomically derived Garway-Heath et al.[45] 2000 map as shown in Figures 4A, 4B, the fit of the HK-LM becomes better ($R^2_{cd} = 0.31$ inferior hemifield, $R^2_{cd} = 0.34$ superior hemifield) while the fit of the H-NLM becomes worse ($R^2_{cd} = -7.17$ inferior hemifield, $R^2_{cd} = -4.22$ superior hemifield).

A Test of the Assumptions Regarding RGC Density as a Function of Eccentricity

The H-NLM depends on estimates of the density of RGC bodies as a function of eccentricity in healthy eyes (Equation S1).[14] Curcio and Allen[54] measured RGC density in postmortem human eyes; their data are plotted in Figure 5A (bold black line) along with a 95% confidence interval for the mean (gray lines) (calculated based on the more conservative *t*-distribution given the relatively small sample size). Superimposed (red line) are the predictions of the study by Harwerth et al.[14] based on monkey data, plotted after adjusting for the difference in the

Investigative Ophthalmology & Visual Science

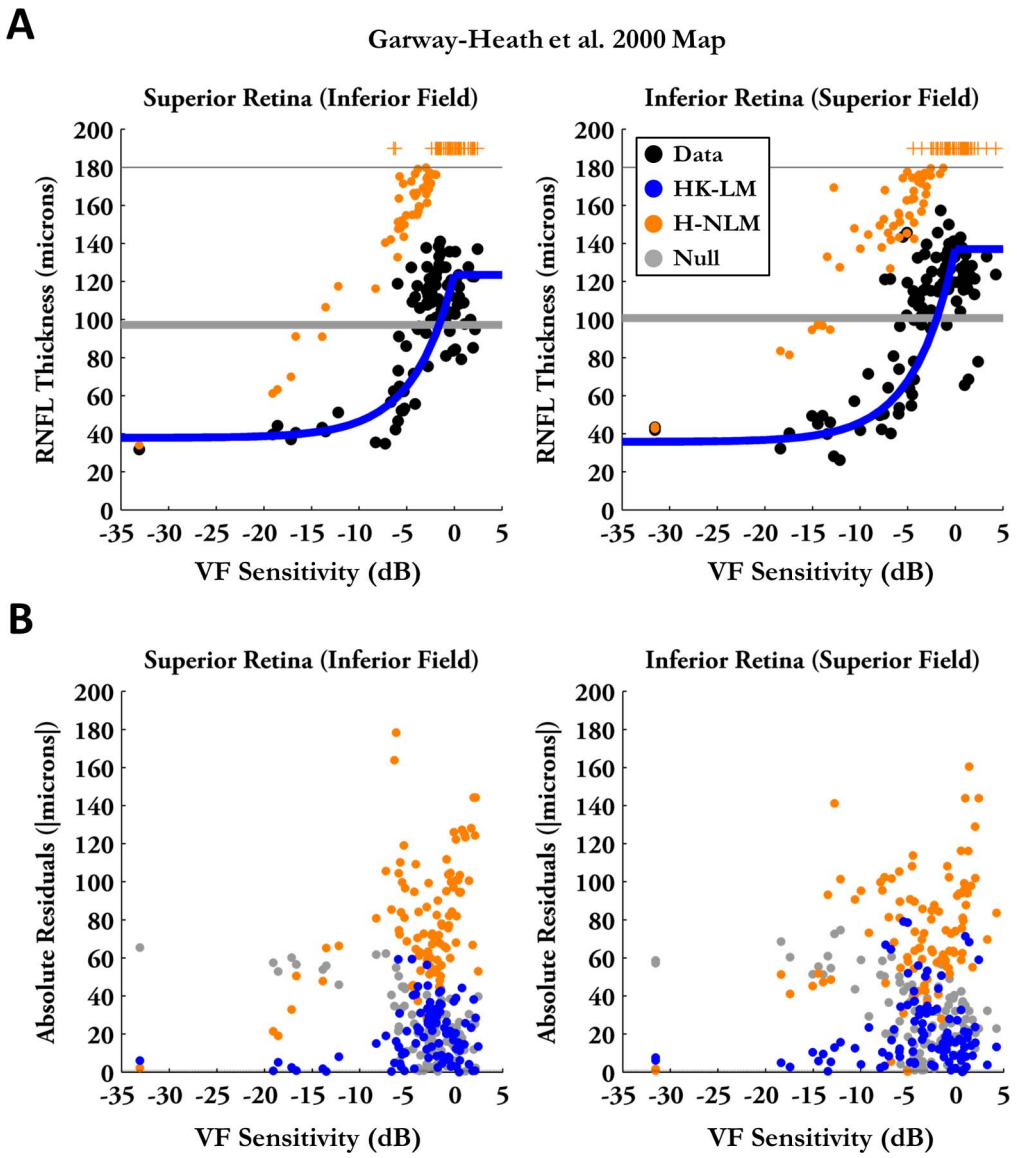

Figure 4. Structure and function, same as in Figure 3, but related using the Garway-Heath et al.[45] 2000 topographic map. The total deviation from age-matched normative data for the VF sensitivity is on the *x*-axis. (**A**) The RNFL thickness for the region corresponding to the VF, based on a topographic map, is on the *y*-axis. The predictions of both the HK-LM (*blue*) and the H-NLM (*orange*) for the data from the glaucoma patients and suspects (*black circles*), along with a "null" model (*gray horizontal line*). Outliers (*plus symbols*) are pinned to the top of the graph. (**B**) The absolute residuals of each of the models from the data (*black circles*) in (**A**).

axial length of the monkey eye. Next, the light sensitivity values for healthy controls were determined via simulation (Methods, Equation 1) for the mean age (35.5 years) of the samples from the Curcio and Allen[54] study, and the density of RGC bodies (green line) was plotted as a function of eccentricity based on the H-NLM (Equations S4, S15, S16, S22) using the latest set of published parameters.

Notably, the predictions of the H-NLM (green line) were markedly above the 95% confidence interval based on the human histology (black line). To better illustrate predictive performance at higher eccentricities, Figure 5B shows the values in Figure 5A scaled relative to the data of Curcio and Allen.[54] Overall, the H-NLM does not predict these data well. Interestingly, a parameter change (in which Equation S3 became S16) in the H-NLM, which also coincided with the newer Harwerth et al.[10] 2010 topographic map, had a very large impact on the relationship between sensitivity and RGC density (Equation S15). The results when using the older parameters (Equation S3) are also plotted (blue line) in Figures 5A, 5B and are closer to the histology data of Curcio and Allen,[54] although notably the estimates for eccentricities near the fovea (less than approximately 4°) and in the periphery (more than approximately 12°) are still outside the 95% confidence intervals. Subsequent work by Harwerth and colleagues,[10,17] as well as by others using their model,[30–33,35] have all used the parameters in Equation S16. Therefore, the parameters from Equation S16 were assumed to be the latest version of the model for further analysis. (Note that most analyses were done with both sets of parameters and the conclusions remained the same.)

Assumptions Regarding Age-Dependent Changes in Structure

Although the original version of the H-NLM did not have an age-dependent correction for structural measures, Harwerth et

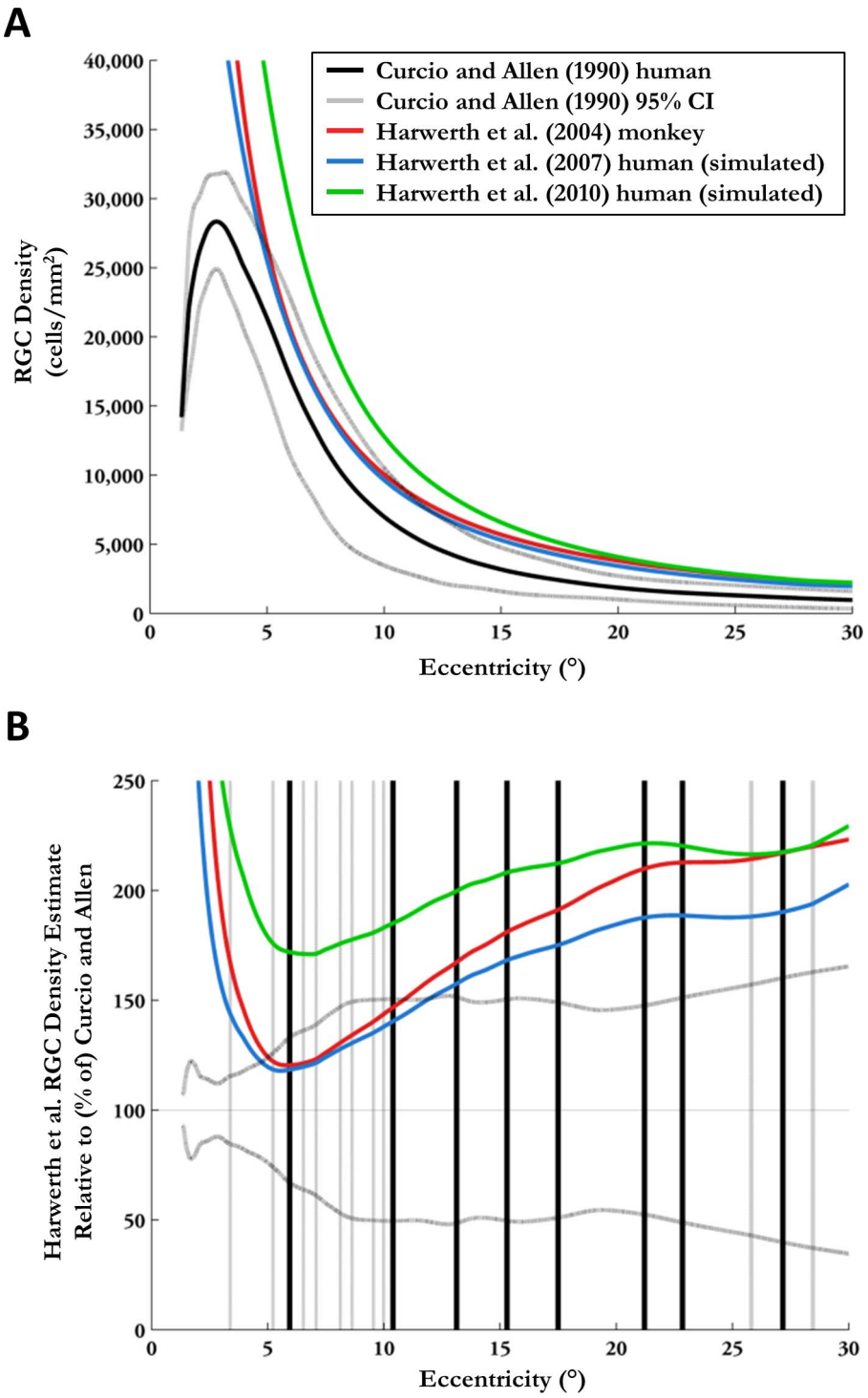

FIGURE 5. (**A**) Retinal ganglion cell density based on the histological measurements (*black*) of Curcio and Allen,[54] with a 95% confidence interval for the mean (*gray lines*), along with predictions of the H-NLM, based on simulated normal sensitivity values (*green*), using an older set of parameters[44] (*blue*), and using the original equations derived based on monkey data[14] (*red*) after correcting for differences in axial length between the monkey and human eye, with details given in the text. (**B**) Same as in (**A**), but expressed as a percentage of the histological measurements (*black line* in [**A**]). The vertical lines indicate the approximate locations of the test stimuli for the 10-2 VF (*light gray*) and the 24-2 VF (*bold black*). Note that the number of samples at each eccentricity (*each vertical line*) varies, which will affect aggregate model error.

al.[15] added the assumption that axon density decreases with age, even in healthy control eyes. Because studies relating RGC-axon count to age using postmortem histology in human subjects have relatively small sample sizes, raw data were digitized from four studies[3,55–57] and obtained directly from data tables provided two others.[58,59] When a study included information from two eyes, the data for the two eyes were averaged. This meta-analysis yielded a sample size of 129 eyes (age = 54.3 ± 22.6 years [mean ± SD]) without optic nerve disease, as shown in Figure 6A. The average axon count for all

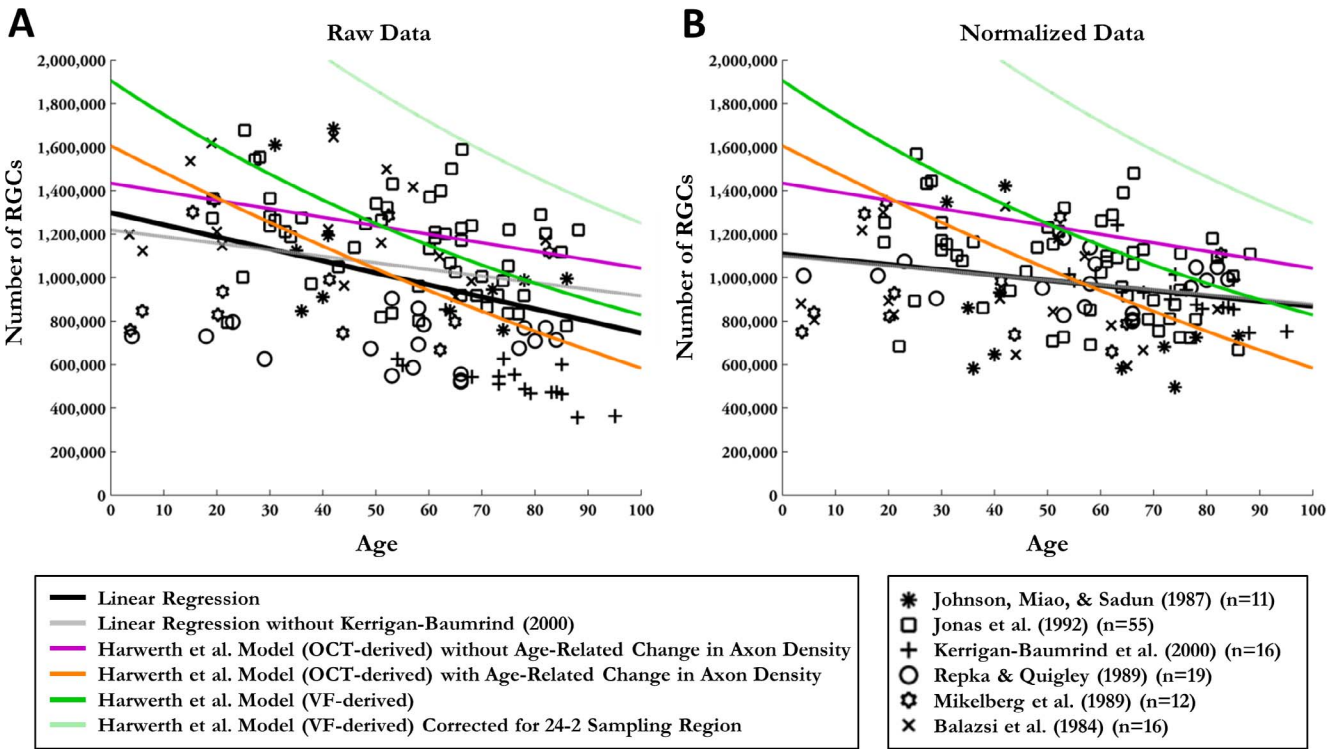

Figure 6. (**A**) The number of RGC axons for postmortem histological measurements of eyes without optic nerve disease from a meta-analysis of the literature[3,55–59] (see legend), along with a best-fit linear regression (*black*), the regression without the Kerrigan-Baumrind et al.[3] study (*gray*), the estimates based on the H-NLM using sensitivity (*green*; corrected for 24-2 VF sampling in *light green*) as well as the estimates based on the H-NLM using OCT without age correction (*magenta*) and with age correction (*orange*). (**B**) Same as in (**A**), but with the literature values normalized across studies (see text for details).

eyes was 998,051 ± 318,830 (mean ± SD). A best-fit linear regression to the histological data (solid black line) yielded a slope of −5535 axons per year ($R^2_P = 0.15$, $P < 0.01$).

Next, the average OCT-derived RNFL thickness for the entire optic disc was simulated based on the linear equation provided by Harwerth et al.[15] (Equation 2), which predicts a decrease in RNFL thickness by approximately 0.3 μm per year. These simulated data were then used as an input into the H-NLM without an age-dependent axon density correction (Equations S8, S11) to determine the effect of age on RGC-axon count for the H-NLM based solely on OCT-derived RNFL thickness (Fig. 6A, magenta line). Notably, although there is a systematic offset to the predictions of the H-NLM that yields a poor fit to the literature data ($R^2_{cd} < 0$), the slope of −3911 axons per year is similar to the line of best-fit (black line). When the age-dependent axon density correction (Equation S12) was implemented into the H-NLM (orange line), the fit was still considerably worse ($R^2_{cd} = 0.02$) than the line of best-fit, and the slope became −10,220 axons per year, almost twice that predicted by the histological data. Next, the change in the number of RGC bodies based on age for the H-NLM (green line) was determined by simulating VF sensitivity data for varying ages, as previously described. Again, the H-NLM function did not fit the data well ($R^2_{cd} < 0$), and predicted a slope of −10,534 axons per year, also almost twice that predicted by the histological data. Interestingly, the addition of the age-dependent correction of the structural data to the H-NLM (change from magenta to orange line) brings the model out of agreement with the histological data from the meta-analysis but brings the two separate RGC estimates (from function and structure) closer together.

Of the data in Figure 6A, a single study by Kerrigan-Baumrind et al.[3] heavily influenced the best-fit line (black). To estimate the effect of possible systematic bias across studies, each study was normalized by addition or subtraction such that the mean of the individuals in the range of 40 to 60 years would be 1 million RGC axons. This normalized meta-analysis is shown in Figure 6B, where the slope of the best-fit line (black) was reduced to −2405 axons per year, whereas the age-adjusted H-NLM (orange) predicts a slope that is more than four times greater.

Assumptions Involved in Evaluations of the H-NLM

The agreement of estimates of RGC bodies and RGC axons (Fig. 2A) has been offered as evidence supporting the H-NLM. In particular, the sensitivity-derived estimates of RGC bodies are plotted against the OCT-derived estimates of RGC axons (or vice versa) and the R^2_P values are reported.[10,44,60,61] Here, a similar graph (Fig. 7) is generated using simulated values based on the H-NLM for healthy controls from ages 20 to 80 years in 5-year increments (solid green circles). In Figure 7, the values on the *y*-axis and *x*-axis were determined based on the orange and green lines, respectively, in Figure 6A. The important point here is that these values can be very well correlated ($R^2_P > 0.99$) even though the predictions of the H-NLM (as in Fig. 6A) do not fit the histological data well (recall that the R^2_{cd} was <0.03 for both the orange and green lines). Thus, the correlation coefficient for plots like that in Figure 7 is better interpreted as a proxy for variability rather than evidence for the validity of the model. Likewise, one could reduce the distance of these predictions (solid green circles in Fig. 7) from unity correlation (thereby reducing the residuals) by increasing the OCT-derived RGC-axon estimates (orange line in Fig. 6A), which would result in an even worse fit to the histological data

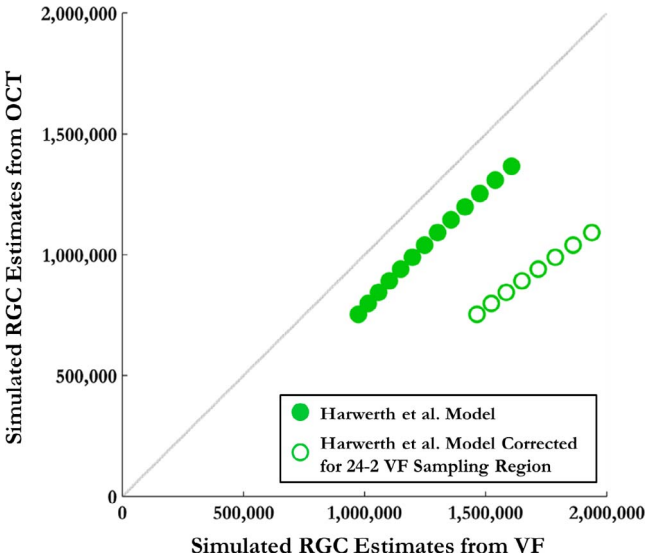

FIGURE 7. The RGC estimates of the H-NLM (*solid green symbols*) using simulated VF data (*x*-axis; *green line* in Fig. 6) and simulated OCT data (*y*-axis; *orange line* in Fig. 6), as well as after adjusting for the 24-2 sampling (*open green symbols*). This figure is plotted in the same form as Figures 2A, 2B, although the axes are in linear units with the same range as the *y*-axis of Figure 6. Note the high correlation between the RGC estimates derived from simulated VF and OCT data despite the differences from histological data in Figure 6.

but a better agreement between the RGC estimates derived separately from function and structure.

Further, when using sensitivity-derived estimates of RGC bodies based on sampling only within approximately ±27° of visual angle from the fovea, the H-NLM should in fact underestimate the true total RGC-axon count for an individual eye. That is, an accurate model, when using the oversimplified map, that compares the 24-2 VF with the entire optic disc, should have points falling above the gray unity line in Figure 7, not close to unity or below unity. Even if the central ±30° of visual angle were sampled by the VF test (as possible with the 30-2 protocol), the remaining region outside 30° corresponds to approximately 73% of the total retinal area, which, even with reduced RGC density, still accounts for approximately 25% of the total RGC count for the average human eye.[54]

To further illustrate this issue, the light sensitivity for the same age range as in Figure 6 was simulated again, this time generating 30-2 VF for healthy controls, which has slightly more extended sampling (out to 30°) than the 24-2. These data were used to generate estimates of RGC bodies based on the H-NLM, and these estimates were further corrected based on the assumption that only approximately 75% of the RGCs were sampled within the central 30°. These new, corrected estimates are shown in Figures 6A, 6B (light green line) and Figure 7 (open green circles). Thus, the predictions of the H-NLM based on sensitivity, once corrected for the region sampled, are markedly higher than the values derived based on structure. These estimates suggest that some of the underlying assumptions of the H-NLM should be revisited.

DISCUSSION

Here, both the H-NLM and HK-LM were compared quantitatively with an independent fdOCT dataset that was not previously used to test or develop either model. In addition, the assumptions underlying the H-NLM were explored. The poor predictive performance of the H-NLM observed with our independent dataset is surprising, as are the observed disagreements between simulations of the H-NLM and human histological data. Below, possible factors contributing to these results are discussed, including the impact of the metric used to assess the fit of the models, changes in particular parameters of the model, assumptions of the model, and iterative changes to the model.

Quantitative Evaluations of the HK-LM and the H-NLM

The performance of the H-NLM in our study was considerably worse than the validation study comparing the H-NLM with the HK-LM performed by Harwerth et al.,[10] which concluded that the performance was similar but that the HK-LM was "less precise" and that the statistical metrics used indicated the "heteroscedasticity" of the HK-LM. However, heteroscedasticity suggests that the variability of errors changes in magnitude across the range of predictions, and this is not an inherent property of the HK-LM any more than it is a property of the H-NLM. Rather, this apparent increase in heteroscedasticity is largely due to differences in the size and location of the region sampled based on the topographic maps. The variability of errors will be larger when a smaller region of the optic disc near the highly vulnerable and thick arcuate regions is sampled (Garway-Heath et al.[45] 2000 map; Fig. 1A) than when data are averaged over a large region of the disc (Harwerth et al.[10] 2010 map and Wheat et al.[17] 2012 map; Figs. 1D, 1E). In fact, quantitative metrics, such as the RMSE, used by Harwerth et al.[10] to compare the H-NLM with the HK-LM, do not take into account the difference in the mean and variability inherent in the datasets themselves due to sampling different regions of the disc when using different topographic maps for each model. When a "null" model (i.e., no change in RNFL thickness with VF sensitivity; gray line in Fig. 3) is evaluated using the RMSE, it appears to perform worse when using the Garway-Heath et al.[45] 2000 map than when using the Harwerth et al.[10] 2010 map or Wheat et al.[17] 2012 map, although its performance is the same across topographic maps when using the R^2_{cd} (Table 1). Thus, the R^2_{cd} should be used to explore the models across a larger variety of maps, although for proper interpretation, models should still be compared pairwise on a specific map, and the anatomical validity of the map itself should be considered. (Note that, given there is variability on both axes, ideally the R^2_{cd} should be calculated again after exchanging the independent and dependent variables. In this study, these values also were determined and a similar trend was observed. However, because this analysis required additional assumptions for both models and because the conclusions remained the same, these values were omitted for brevity and simplicity.) As seen in Table 2, the performance of the HK-LM decreases when the topographic maps are not accurate with regard to the spatial correspondence between the VF and OCT sampling regions (e.g., Figs. 1C–F), whereas the performance of the H-NLM increases.

Even using the Harwerth et al.[10] 2010 map, the H-NLM R^2_{cd} values remain negative in our study as compared with the reported values of $R^2_{cd} = 0.4$ (both hemifields combined) in the validation study by Harwerth et al.[10] Because the most recently published parameters from the Wheat et al.[17] 2012 study differ slightly from those used in the Harwerth et al.[10] 2010 study, our data were reanalyzed using the parameters from the Harwerth et al.[10] 2010 study. However, the performance of the H-NLM remained poor when using either the Garway-Heath et al.[45] 2000 map ($R^2_{cd} = -6.88$ inferior hemifield, $R^2_{cd} = -3.94$ superior hemifield) or the Harwerth et al.[10] 2010 map ($R^2_{cd} = -6.15$ inferior hemifield, $R^2_{cd} = -3.17$ superior hemifield), although the values were slightly less negative. In fact, various

Table 1. The Effect of Different Topographic Maps on Metrics of Evaluation

	R^2_{cd}			RMSE		
	HK-LM	H-NLM	Null	HK-LM	H-NLM	Null
Map: Garway-Heath et al.[45] 2000						
Superior retina (inferior field)	0.31	−7.17	0	24	84	30
Inferior retina (superior field)	0.34	−4.22	0	28	80	35
Map: Harwerth et al.[10] 2010						
Superior retina (inferior field)	0.15	−6.86	0	18	56	20
Inferior retina (superior field)	0.12	−3.46	0	19	43	20
Map: Wheat et al.[17] 2012						
Superior retina (inferior field)	0.15	−6.86	0	18	56	20
Inferior retina (superior field)	0.14	−3.18	0	19	42	20

Note that the null model appears to perform worse on the Garway-Heath et al.[45] 2000 map when using the RMSE.

single-parameter changes from the most recently published set of parameters in the Wheat et al.[17] study were considered, and none produced an R^2_{cd} value above zero (see Supplementary Table S1). In agreement with the analysis shown in Figure 5, the most beneficial change was obtained by altering the value of 0.9 in Equation S16 (green line in Fig. 5) to a previously used value of 0.95, as in Equation S3 (blue line in Fig. 5).

We were able to further improve the fit of the H-NLM by both using the value of 0.95 from Equation S3 and by removing the disease-stage–dependent correction, the most recent addition to the H-NLM. Although these steps resulted in an overall improved performance of the H-NLM with the Garway-Heath et al.[45] 2000 map (R^2_{cd} = −0.41 inferior hemifield, R^2_{cd} = −0.11 superior hemifield), the R^2_{cd} remained less than zero, and these changes increased the disagreement between RGC estimates (Fig. 2B). Given that the Garway-Heath et al.[45] 2000 map is more anatomically accurate, this calls into question the justification for the disease-stage–dependent correction, as well as the accuracy of the RGC estimates. Also, note that although Equation S3 yielded an improved fit, it is Equation S16 that is used in the Harwerth et al.[10] validation study and subsequent publications,[17,30–33,35,60] including those using the H-NLM as a basis for further modeling. Finally, because removing the disease-stage–dependent correction appeared to be beneficial, we explored the effect of testing the entire range of data (including controls and ION patients) while maintaining the favorable value of 0.95 from Equation S3. Under these conditions, the fit of the H-NLM with the Garway-Heath et al.[45] 2000 map (R^2_{cd} = −0.76 inferior hemifield, R^2_{cd} = −0.19 superior hemifield) nonetheless improved when removing the disease-stage–dependent correction (R^2_{cd} = −0.05 inferior hemifield, R^2_{cd} = 0.22 superior hemifield). Although this analysis yielded a positive R^2_{cd} value for the superior hemifield, recall that for the entire range of data, using previously published parameters based on tdOCT, the HK-LM performance on the Garway-Heath et al.[45] 2000 map was R^2_{cd} = 0.64 (inferior hemifield) and R^2_{cd} = 0.52 (superior hemifield), considerably better than the best performance of the H-NLM even after exploring different combinations of parameters. Also, note that using the entire range of data for the H-NLM with previously published parameter values from either the Wheat et al.[17] 2012 study or the Harwerth et al.[10] 2010 study did not result in any R^2_{cd} values greater than zero irrespective of the topographic map used. Thus, the major issues concerning H-NLM performance raised here remain pertinent, regardless of which parameters are used for the H-NLM. Furthermore, the relatively strong performance of the HK-LM using parameters based on the tdOCT, as well as the general agreement between tdOCT and fdOCT,[62] should preclude any arguments regarding the need for any additional parameter to adjust for the fdOCT, as has been previously suggested.[10]

Table 2. The R^2_{cd} values for both the HK-LM and the H-NLM Across Several Different Topographic Maps

	R^2_{cd}	
	HK-LM	H-NLM
Map: Garway-Heath et al.[45] 2000		
Superior retina (inferior field)	0.31	−7.17
Inferior retina (superior field)	0.34	−4.22
Map: Garway-Heath et al.[46] 2002		
Superior retina (inferior field)	0.34	−9.12
Inferior retina (superior field)	0.35	−12.58
Map: Harwerth et al.[44] 2007		
Superior retina (inferior field)	0.16	−2.71
Inferior retina (superior field)	0.17	−0.37
Map: Harwerth et al.[10] 2010		
Superior retina (inferior field)	0.15	−6.86
Inferior retina (superior field)	0.12	−3.46
Map: Wheat et al.[17] 2012		
Superior retina (inferior field)	0.15	−6.86
Inferior retina (superior field)	0.14	−3.18
Map: Medeiros et al.[60] 2012		
Superior retina (inferior field)	0.16	−3.29
Inferior retina (superior field)	0.02	−1.49
Map: oversimplified		
Entire retina (∼±27° field)	0.11	−2.55

Evaluating the Assumptions of the H-NLM

First, the topographic maps used in studies of the H-NLM do not agree well with some aspects of the anatomy. In particular, the Harwerth et al.[44] 2007 map (Fig. 1C), particularly sectors 3 and 8, does not seem to agree with the path of RGC axons (Fig. 1G, middle), and the other topographic maps (Figs. 1D–F) include large portions of the nasal half of the optic disc, which receive input from outside the region sampled by the 24-2 VF (Fig. 1G, middle). It should be noted that the concerns regarding these maps are true independent of the model in which they are used, although they are particularly problematic when comparing estimates of RGCs derived from measures of structure and function. Moreover, when using an oversim-

Investigative Ophthalmology & Visual Science

plified topographic map to compare RGC estimates for the entire eye, one would in fact expect disagreement for an accurate model (not agreement as in Figs. 2A, 7). Second, there are problems with the assumptions involved in the RGC estimates. One of the fundamental equations of the H-NLM (Equation S15) yields predictions for RGC density that differ markedly from the histological data in human eyes (Fig. 5). A more general point is that many of the original parameter estimates of the H-NLM (e.g., Equation S2) were derived using ordinary least-squares regression lines, which is problematic when there is measurement error on both axes. Third, the agreement between two independently derived estimates for the RGC count is compelling evidence for the validity of a model estimating RGCs only if both estimates are entirely without empirically fit parameters. This is a key point. Once any model is revised to improve the agreement between the two measures, such agreement is no longer compelling evidence that the model accurately describes histology. In fact, continuous adjustments to maintain this agreement may cause the model to deviate further from an accurate estimation.

Evaluating the Value-Added of Assumptions That Increase Model Complexity

Moreover, it is not clear if the additional assumptions of the H-NLM, which increase the complexity of the model, yield substantial value-added. For instance, the model appears to overestimate the effect of age on the number of RGC axons in healthy controls (Figs. 6A, 6B). The amount of axon loss estimated by the model (more than 10,000 axons per year; i.e., more than 100,000 axons per decade) is between two and four times greater than that derived from a meta-analysis of the data available in the literature for human histology. Note that a loss of more than 100,000 axons per decade at a linear rate implies that an individual starting life with 800,000 axons would have none left by 80 years of age. Although data from monkeys was ignored in the meta-analysis shown here, a recent study by Fortune et al.[63] is noteworthy as, in 46 monkeys, approximately 100% of the optic nerve was sampled for each animal using an automated and validated counting technique.[64] The estimate of axon loss per year from Fortune et al.,[63] after taking into consideration the shorter life span of the monkey, is quite close to the estimate of between 2200 and 3900 axons per year derived from the meta-analysis done here. A general point is that the addition of age correction in the H-NLM increased the agreement between sensitivity-derived and OCT-derived estimates but may have decreased the correspondence to histological data. For a model that estimates RGCs, it is particularly important to realize that increasing agreement between these two estimates, for example, by introducing an age correction into the OCT-derived estimated RGC-axon count, assumes that the other estimate, the sensitivity-derived RGC body count, is accurate. An alternate explanation, which should be considered in the future, is simply that some of the previous assumptions require reconsideration without introducing further complexity. For example, as previously discussed by Fortune et al.,[63] the fact that sensitivity is not age-corrected in the H-NLM may lead to an overestimation of the effect of age on the number of RGC bodies; that is, perhaps part of the decline of sensitivity with age is due to factors not directly related to RGC density.

A different aspect of the H-NLM that increases complexity without clear value-added is the assumption regarding disease-stage–dependent changes in structure. In patients with extreme losses in VF sensitivity, the OCT-measured RNFL has a residual thickness.[11,21] The earlier versions[14,44] of the H-NLM did not account for this residual, assuming that RNFL thickness eventually decreased to 0 μm with severe visual loss. In contrast, the HK-LM assumes that a residual thickness, composed of glial cells and blood vessels, is reached. The H-NLM was later modified[10,17] to include a residual component. Further, the newer version of the H-NLM assumes that gliosis increases with disease stage, thus increasing the residual thickness and altering the nature of the structure–function relationship. Although the correction factor in Equation S18 as defined in Equation S23 may seem like an independent correction for the stage of disease, in fact the inclusion of the average total deviation as a parameter in Equation S23 is equivalent to adding another sensitivity-dependent term, fundamentally altering the underlying relationship between sensitivity and structure. It is noteworthy that both Equations S20 and S23 have a nonzero intercept, such that a value of deviation $D = 0$, as expected for normal sensitivity, still yields a correction factor for the overall estimated RGC-axon count. Similarly, if gliosis is the explanation for this correction factor, one would expect the function to be bounded such that it would not be applicable for above-normal sensitivity values; however, this is not the case. Finally, recall that removing this assumption from the H-NLM improved the fit of the model to our data, although at the cost of decreasing agreement (Fig. 2B) between RGC estimates derived from structure and function.

Recommendations and Future Directions

Several of the fundamental assumptions of the H-NLM need to be reconsidered. In general, our approach here has erred on the side of being conservative. For instance, we ignored the difference in the number of parameters between the HK-LM and H-NLM, using the R^2_{cd} rather than the adjusted R^2_{cd}, which penalizes a model for the complexity of additional parameters. However, future work also should consider the relative complexity of each model when comparing their performance. Admittedly, the fit of the HK-LM may improve with additional parameters. For instance, although it uses age-corrected VF sensitivity values, the HK-LM assumes that the change in RNFL thickness with age is relatively small and currently does not include a correction for the OCT data. However, we have previously shown[22] that, given the variability inherent in each measure, it is difficult to assess the nature of the structure–function relationship in glaucoma. Therefore, the evidence for a more complex model should be compelling.

Moreover, it is probably advisable to separate the issue of estimating RGCs from examining the structure–function relationship. Although it is understandably attractive to be able to express the structure–function relationship in terms of RGCs,[61] particularly with regard to glaucomatous changes, it is possible to increase the structure–function agreement while decreasing the accuracy of RGC estimation. However, we do not mean to suggest that the previous work by Harwerth et al.[9,10] has not been worthwhile; among other advantages, it has certainly sharpened the debate in the field regarding the assumptions underlying structure–function models. Nonetheless, particularly given the importance of the past work done by Harwerth et al. (e.g., Ref. 14) using histology, moving forward, it would be preferable to return to such histological measures as a method of validating and further refining these RGC estimates.

Likewise, we do not mean to suggest that all work building on the H-NLM is invalid. For instance, the general notion behind the Medeiros et al.[30] combined structure–function index (CSFI), which seems to be geared toward progression, is that the dynamic range is greater for sensitivity loss when disease becomes severe. Although it is important to keep in mind the recent concerns raised by Gardiner et al.[65] regarding the accuracy of reported sensitivity losses beyond approximately 15 to 19 dB, certainly it seems plausible that the OCT

Investigative Ophthalmology & Visual Science

may hit a noise floor before the VF for advanced disease. That said, the fact that models building on the H-NLM, such as CSFI, are able to estimate RGC counts should not be used to argue that these models are inherently better. Such models should be evaluated empirically and not treated as superior because of the use of an RGC estimate, particularly given the increased complexity of the model. In fact, given the increased complexity, it may be worthwhile to explore the performance of the CSFI without estimating RGCs to determine the extent of value-added. In any case, more work needs to be done before recommending the use of the CSFI in clinical trials, particularly if part of the basis for that claim relies on the relationship of the CSFI to estimated RGC counts.

However, other than the work regarding the CSFI by Medeiros et al.,[30] there have been other publications that have used the H-NLM to estimate RGCs where it is difficult to argue that the putative value-added is not tied closely to the validity of the RGC estimates. For instance, the recent work of Marvasti et al.[36] argues that the VF index, used to monitor progression, underestimates the amount of neural loss in glaucoma. In a similar vein, Tatham et al.[33] argues that even relatively local RNFL defects can be associated with large neuronal losses in glaucoma. Taken independently, it is certainly reasonable that the VF index may underestimate progression or that considerable neuronal loss may be associated with early, local defects. However, it is difficult to disassociate the impact of these studies from their claims regarding RGC estimates, and therefore the importance of such studies is more closely tied to the validity of the H-NLM RGC estimates.

Finally, given the increasing potential of newer OCT technology, it may be possible to estimate the number of RGCs directly from RGC thickness measurements. This may provide a method of estimating RGCs with a considerably smaller set of assumptions. Additionally, independent of RGC estimates, further work is needed to develop a priori predictive models that relate OCT-derived RGC thickness measurements to visual sensitivity.

Conclusions

Here we examined the assumptions underlying the increasingly prominent model of Harwerth et al.,[10] which estimates the number of RGCs from both visual sensitivity and OCT-derived RNFL thickness measurements. Our results indicate that the predictions of the Harwerth et al.[10] model differ noticeably from histological data and several of the assumptions underlying the model need to be reexamined, including the topographic map. Further, when the Harwerth et al.[10] model was applied to an independent dataset, it performed poorly, whereas a simple linear model by Hood and Kardon[11] performed reasonably well. Thus, the approach of estimating RGCs and the resulting complexity of the Harwerth et al.[10] model does not appear to yield value-added when compared with the simpler approach of Hood and Kardon.[11] In any case, studies and models relying on the RGC estimates of the H-NLM should be interpreted with caution.

Acknowledgments

The authors thank Gustavo De Moraes, MD, Jeffrey Liebmann, MD, Robert Ritch, MD, and Jeffrey Odel, MD, for patient referrals; Randy Kardon, MD, PhD, for helpful discussions; Norma Graham, PhD, for helpful discussions and for providing critical revisions to the text; Alyssa Ehrlich for providing critical revisions to the text; Daiyan Xin for assisting with data acquisition; and David Rhee and Diane Wang for helping with the manual correction of the retinal layer segmentation from the OCT images.

Supported by National Institutes of Health Grant R01-EY-02115 (DCH) and National Science Foundation Graduate Research Fellowship Grant DGE-11-44155 (ASR).

Disclosure: **A.S. Raza**, None; **D.C. Hood**, Topcon, Inc. (F)

References

1. Quigley HA, Addicks EM, Green WR. Optic nerve damage in human glaucoma: III. Quantitative correlation of nerve fiber loss and visual field defect in glaucoma, ischemic neuropathy, papilledema, and toxic neuropathy. *Arch Ophthalmol*. 1982; 100:135–146.
2. Quigley HA, Dunkelberger GR, Green WR. Retinal ganglion cell atrophy correlated with automated perimetry in human eyes with glaucoma. *Am J Ophthalmol*. 1989;107:453–464.
3. Kerrigan-Baumrind LA, Quigley HA, Pease ME, Kerrigan DF, Mitchell RS. Number of ganglion cells in glaucoma eyes compared with threshold visual field tests in the same persons. *Invest Ophthalmol Vis Sci*. 2000;41:741–748.
4. Huang D, Swanson EA, Lin CP, et al. Optical coherence tomography. *Science*. 1991;254:1178–1181.
5. Sharma P, Sample PA, Zangwill LM, Schuman JS. Diagnostic tools for glaucoma detection and management. *Surv Ophthalmol*. 2008;53:17–32.
6. Chang R, Budenz DL. New developments in optical coherence tomography for glaucoma. *Curr Opin Ophthalmol*. 2008;19: 127–135.
7. Curcio C, Messinger JD, Sloan KR, Mitra A, McGwin G, Spaide RF. Human chorioretinal layer thicknesses measured in macula-wide, high-resolution histologic sections. *Invest Ophthalmol Vis Sci*. 2011;52:3943–3954.
8. Spaide R, Curcio C. Anatomical correlates to the bands seen in the outer retina by optical coherence tomography. *Retina*. 2011;31:1609–1619.
9. Harwerth RS, Crawford MLJ, Frishman LJ, Viswanathan S, Smith EL, Carter-Dawson L. Visual field defects and neural losses from experimental glaucoma. *Prog Retin Eye Res*. 2002; 21:91–125.
10. Harwerth RS, Wheat JL, Fredette MJ, Anderson DR. Linking structure and function in glaucoma. *Prog Retin Eye Res*. 2010; 29:249–271.
11. Hood DC, Kardon RH. A framework for comparing structural and functional measures of glaucomatous damage. *Prog Retin Eye Res*. 2007;26:688–710.
12. Wollstein G, Kagemann L, Bilonick RA, et al. Retinal nerve fibre layer and visual function loss in glaucoma: the tipping point. *Br J Ophthalmol*. 2012;96:47–52.
13. Malik R, Swanson WH, Garway-Heath DF. structure–function relationship in glaucoma: past thinking and current concepts. *Clin Experiment Ophthalmol*. 2012;40:369–380.
14. Harwerth RS, Carter-Dawson L, Smith EL, Barnes G, Holt WF, Crawford MLJ. Neural losses correlated with visual losses in clinical perimetry. *Invest Ophthalmol Vis Sci*. 2004;45:3152–3160.
15. Harwerth RS, Wheat JL, Rangaswamy N. Age-related losses of retinal ganglion cells and axons. *Invest Ophthalmol Vis Sci*. 2008;49:4437–4443.
16. Harwerth RS, Wheat JL. Modeling the effects of aging on retinal ganglion cell density and nerve fiber layer thickness. *Graefes Arch Clin Exp Ophthalmol*. 2008;246:305–314.
17. Wheat JL, Rangaswamy N, Harwerth RS. Correlating RNFL thickness by OCT with perimetric sensitivity in glaucoma patients. *J Glaucoma*. 2012;21:95–101.
18. Hood DC, Greenstein V, Odel J, et al. Visual field defects and multifocal visual evoked potentials: evidence of a linear relationship. *Arch Ophthalmol*. 2002;120:1672–1681.

19. Hood DC, Anderson SC, Wall M, Kardon RH. Structure versus function in glaucoma: an application of a linear model. *Invest Ophthalmol Vis Sci*. 2007;48:3662–3668.
20. Hood DC. Relating retinal nerve fiber thickness to behavioral sensitivity in patients with glaucoma: application of a linear model. *J Opt Soc Am A Opt Image Sci Vis*. 2007;24:1426–1430.
21. Hood DC, Anderson S, Rouleau J, et al. Retinal nerve fiber structure versus visual field function in patients with ischemic optic neuropathy. A test of a linear model. *Ophthalmology*. 2008;115:904–910.
22. Hood DC, Anderson SC, Wall M, Raza AS, Kardon RH. A test of a linear model of glaucomatous structure–function loss reveals sources of variability in retinal nerve fiber and visual field measurements. *Invest Ophthalmol Vis Sci*. 2009;50:4254–4266.
23. Wojtkowski M, Fercher AF, Leitgeb R. Phase-sensitive interferometry in optical coherence tomography. *Proc SPIE*. 2001; 4515:250–255.
24. Vizzeri G, Bowd C, Medeiros FA, Weinreb RN, Zangwill LM. Effect of improper scan alignment on retinal nerve fiber layer thickness measurements using stratus optical coherence tomograph. *J Glaucoma*. 2008;17:341–349.
25. Gabriele ML, Ishikawa H, Wollstein G, et al. Optical coherence tomography scan circle location and mean retinal nerve fiber layer measurement variability. *Invest Ophthalmol Vis Sci*. 2008;49:2315–2321.
26. Horn FK, Mardin CY, Laemmer R, et al. Correlation between local glaucomatous visual field defects and loss of nerve fiber layer thickness measured with polarimetry and spectral domain OCT. *Invest Ophthalmol Vis Sci*. 2009;50:1971–1977.
27. Rao H, Zangwill LM, Weinreb RN, Leite MT, Sample PA, Medeiros FA. structure–function relationship in glaucoma using spectral-domain optical coherence tomography. *Arch Ophthalmol*. 2011;129:864–871.
28. Leite MT, Zangwill LM, Weinreb RN, Rao HL, Alencar LM, Medeiros FA. structure–function relationships using the cirrus spectral domain optical coherence tomograph and standard automated perimetry. *J Glaucoma*. 2012;21:49–54.
29. Pinto LM, Costa EF, Melo LAS, et al. structure–function correlations in glaucoma using matrix and standard automated perimetry versus time-domain and spectral-domain OCT devices. *Invest Ophthalmol Vis Sci*. 2014;55:3074–3080.
30. Medeiros FA, Lisboa R, Weinreb RN, Girkin CA, Liebmann JM, Zangwill LM. A combined index of structure and function for staging glaucomatous damage. *Arch Ophthalmol*. 2012;130: 1107.
31. Medeiros FA, Lisboa R, Weinreb RN, Liebmann JM, Girkin CA, Zangwill LM. Retinal ganglion cell count estimates associated with early development of visual field defects in glaucoma. *Ophthalmology*. 2012;120:1–9.
32. Tatham AJ, Weinreb RN, Zangwill LM, Liebmann JM, Girkin CA, Medeiros FA. The relationship between cup-to-disc ratio and estimated number of retinal ganglion cells. *Invest Ophthalmol Vis Sci*. 2013;54:3205–3214.
33. Tatham AJ, Weinreb RN, Zangwill LM, Liebmann JM, Girkin CA, Medeiros FA. Estimated retinal ganglion cell counts in glaucomatous eyes with localized retinal nerve fiber layer defects. *Am J Ophthalmol*. 2013;156:578–587.e1.
34. Tatham AJ, Meira-Freitas D, Weinreb RN, Marvasti AH, Zangwill LM, Medeiros FA. Estimation of retinal ganglion cell loss in glaucomatous eyes with a relative afferent pupillary defect. *Invest Ophthalmol Vis Sci*. 2014;55:513–522.
35. Meira-Freitas D, Lisboa R, Tatham AJ, et al. Predicting progression in glaucoma suspects with longitudinal estimates of retinal ganglion cell counts. *Invest Ophthalmol Vis Sci*. 2013;54:4174–4183.
36. Marvasti AH, Tatham AJ, Zangwill LM, et al. The relationship between visual field index and estimated number of retinal ganglion cells in glaucoma. *PLoS One*. 2013;8:e76590.
37. Tatham AJ, Weinreb RN, Medeiros FA. Strategies for improving early detection of glaucoma: the combined structure–function index. *Clin Ophthalmol*. 2014;8:611–621.
38. Lisboa R, Weinreb RN, Medeiros F. Combining structure and function to evaluate glaucomatous progression: implications for the design of clinical trials. *Curr Opin Pharmacol*. 2013; 13:115–122.
39. Hood DC, Raza AS, De Moraes CGV, Johnson CA, Liebmann JM, Ritch R. The nature of macular damage in glaucoma as revealed by averaging optical coherence tomography data. *Transl Vis Sci Technol*. 2012;1:1–15.
40. Budenz DL, Rhee P, Feuer WJ, McSoley J, Johnson CA, Anderson DR. Sensitivity and specificity of the swedish interactive threshold algorithm for glaucomatous visual field defects. *Ophthalmology*. 2002;109:1052–1058.
41. Yang Q, Reisman CA, Wang Z, et al. Automated layer segmentation of macular OCT images using dual-scale gradient information. *Opt Express*. 2010;18:21293–21307.
42. Hood DC, Cho J, Raza AS, Dale EA, Wang M. Reliability of a computer-aided manual procedure for segmenting optical coherence tomography scans. *Optom Vis Sci*. 2011;88:113.
43. Raza AS, Cho J, De Moraes CGV, et al. Retinal ganglion cell layer thickness and local visual field sensitivity in glaucoma. *Arch Ophthalmol*. 2011;129:1529–1536.
44. Harwerth RS, Vilupuru AS, Rangaswamy N, Smith EL. The relationship between nerve fiber layer and perimetry measurements. *Invest Ophthalmol Vis Sci*. 2007;48:763–773.
45. Garway-Heath DF, Poinoosawmy D, Fitzke FW, Hitchings RA. Mapping the visual field to the optic disc in normal tension glaucoma eyes. *Ophthalmology*. 2000;107:1809–1815.
46. Garway-Heath DF, Holder GE, Fitzke FW, Hitchings RA. Relationship between electrophysiological, psychophysical, and anatomical measurements in glaucoma. *Invest Ophthalmol Vis Sci*. 2002;43:2213–2220.
47. Jansonius NM, Nevalainen J, Selig B, et al. A mathematical description of nerve fiber bundle trajectories and their variability in the human retina. *Vision Res*. 2009;49:2157–2163.
48. Denniss J, McKendrick AM, Turpin A. An anatomically customizable computational model relating the visual field to the optic nerve head in individual eyes. *Invest Ophthalmol Vis Sci*. 2012;53:6981–6990.
49. Carreras FJ, Medina J, Ruiz-Lozano M, Carreras I, Castro JL. Virtual tissue engineering and optic pathways: plotting the course of the axons in the retinal nerve fiber layer. *Invest Ophthalmol Vis Sci*. 2014;55:3107–3119.
50. Airaksinen PJ, Doro S, Veijola J. Conformal geometry of the retinal nerve fiber layer. *Proc Natl Acad Sci*. 2008;105:19690–19695.
51. Heijl A, Lindgren G, Olsson J. Normal variability of static perimetric threshold values across the central visual field. *Arch Ophthalmol*. 1987;105:1544–1549.
52. *Humphrey Field Analyzer User Manual*. Dublin, CA: Carl Zeiss Meditec, Inc.; 2012:K3–K7.
53. Budenz DL, Anderson DR, Varma R, et al. Determinants of normal retinal nerve fiber layer thickness measured by stratus OCT. *Ophthalmology*. 2007;114:1046–1052.
54. Curcio C, Allen K. Topography of ganglion cells in human retina. *J Comp Neurol*. 1990;300:5–25.
55. Balazsi A, Rootman J, Drance S, Schulzer M, Douglas G. The effect of age on the nerve fiber population of the human optic nerve. *Am J Ophthalmol*. 1984;97:760–766.

Investigative Ophthalmology & Visual Science

56. Jonas JB, Schmidt AM, Müller-Bergh JA, Schlötzer-Schrehardt UM, Naumann GO. Human optic nerve fiber count and optic disc size. *Invest Ophthalmol Vis Sci*. 1992;33:2012–2018.
57. Mikelberg FS, Drance SM, Schulzer M, Yidegiligne HM, Weis MM. The normal human optic nerve. *Ophthalmology*. 1989; 96:1325–1328.
58. Johnson BM, Miao M, Sadun AA. Age-related decline of human optic nerve axon populations. *Age*. 1987;10:5–9.
59. Repka MX, Quigley HA. The effect of age on normal human optic nerve fiver number and diameter. *Ophthalmology*. 1989; 96:26–32.
60. Medeiros FA, Zangwill LM, Anderson DR, et al. Estimating the rate of retinal ganglion cell loss in glaucoma. *Am J Ophthalmol*. 2012;154:814–824.
61. Harwerth RS. A neuron doctrine for glaucoma. *Optom Vis Sci*. 2008;85:436–444.
62. Hood DC, Raza AS, Kay KY, et al. A comparison of retinal nerve fiber layer (RNFL) thickness obtained with frequency and time domain optical coherence tomography (OCT). *Opt Express*. 2009;17:3997.
63. Fortune B, Reynaud J, Cull G, Burgoyne CF, Wang L. The effect of age on optic nerve axon counts, SDOCT scan quality, and peripapillary retinal nerve fiber layer thickness measurements in rhesus monkeys. *Transl Vis Sci Technol*. 2014;3:2.
64. Reynaud J, Cull G, Wang L, et al. Automated quantification of optic nerve axons in primate glaucomatous and normal eyes—method and comparison to semi-automated manual quantification. *Invest Ophthalmol Vis Sci*. 2012;53:2951–2959.
65. Gardiner SK, Swanson WH, Goren D, Mansberger SL, Demirel S. Assessment of the reliability of standard automated perimetry in regions of glaucomatous damage. *Ophthalmology*. 2014;121:1359–1369.

Investigative Ophthalmology & Visual Science