HHS Public Access

Author manuscript

Ophthalmol Glaucoma. Author manuscript; available in PMC 2023 September 01.

Published in final edited form as:

Ophthalmol Glaucoma. 2022 ; 5(5): 507–515. doi:10.1016/j.ogla.2022.02.001.

Clinicians' Use of Quantitative Information when Assessing the Rate of Structural Progression in Glaucoma

Stuart K Gardiner[1,*], **Robert M Kinast**[1], **Teresa C. Chen**[2], **Nicholas G Strouthidis**[3,4], **Carlos Gustavo De Moraes**[5], **Kouros Nouri-Mahdavi**[6], **Jonathan S Myers**[7], **Jin Wook Jeoung**[8], **John T Lind**[9], **Lindsay A. Rhodes**[10], **Donald L Budenz**[11], **Steven L Mansberger**[1]

[1]Devers Eye Institute, Legacy Health, Portland, OR, USA

[2]Harvard Medical School, Massachusetts Eye & Ear, Boston, MA, USA

[3]NIHR Biomedical Research Centre, Moorfields Eye Hospital NHS Foundation Trust and UCL Institute of Ophthalmology, London, UK

[4]Discipline of Clinical Ophthalmology and Eye Health, University of Sydney, Sydney, NSW, Australia

[5]Columbia University Irving Medical Center, New York, NY, USA

[6]Stein Eye Institute, University of California Los Angeles, Los Angeles, CA, USA

[7]Wills Eye Hospital, Philadelphia, PA, USA

[8]Seoul National University Hospital, Seoul National University College of Medicine, Seoul, South Korea

[9]Glick Eye Institute, Indiana University School of Medicine, Indianapolis, IN USA

[10]University of Alabama at Birmingham, Birmingham, AL, USA

[11]Department of Ophthalmology, University of North Carolina-Chapel Hill, Chapel Hill, NC, USA

Abstract

Purpose—OCT scans contain large amounts of information, but clinicians often rely on reported layer thicknesses when assessing the rate of glaucomatous progression. We sought to determine which of these quantifications most closely relate to the subjective assessment of glaucoma experts who had all the diagnostic information available.

Design—Prospective cohort study.

[*]Corresponding Author: Dr. Stuart Gardiner, Devers Eye Institute, Legacy Research Institute, 1225 NE 2nd Ave, Portland, OR 97232, USA, Tel: +1 503 413 1199, sgardiner@deverseye.org.

Meeting Presentation: Presented in part at the 2021 ARVO Meeting.

This article contains additional online-only material. Supplemental Figure 1 should appear online-only.

Participants—Eleven glaucoma specialists independently scored the rate of structural progression from series of 5 biannual clinical OCT printouts.

Methods—100 glaucoma or glaucoma suspect eyes of 51 participants were included; 20 were scored twice to assess repeatability. Scores ranged from 1 (improvement) to 7 (very rapid progression). Generalized estimating equation linear models were used to predict the mean clinician score from the rates of change of Retinal Nerve Fiber Layer Thickness (RNFLT) or Minimum Rim Width (MRW), either globally or in the most rapidly thinning of six sectors.

Main Outcome Measures—The correlation between objective rates of change and the average of the 11 clinicians' scores.

Results—Average RNFLT within series of study eyes was 79.3μm (range: 41.4 to 126.6). 95% of individual clinician scores varied by ≤1 point when repeated. The mean clinician score was more strongly correlated with the rate of change of RNFLT in the most rapidly changing sector in %/y (pseudo-R^2=0.657) than the rate of global RNFLT (0.372). The rate of MRW in the most rapidly changing sector had pseudo-R^2=0.149.

Conclusions—The rate of change of RNFLT in the most rapidly changing sector predicted experts' assessment of the rate of structural progression better than global rates or MRW. Sectoral rates may be a useful addition to current clinical printouts.

Précis

Experienced clinicians subjectively assessed the rate of glaucomatous progression in 100 eyes. Their scores correlated best with the rate of change in the more rapidly thinning sector of the retinal nerve fiber layer.

When assessing the rate of progression in eyes with glaucoma, visual function is of paramount importance to the patient, whereas current structural measures are surrogates for cell loss whose impact on daily living is both indirect and highly variable between patients.[1–4] Yet, structural testing does provide a great amount of additional useful information.[5] Its typically lower test-retest variability[6] might imbue OCT with quicker detection and more accurate prognostic ability,[7, 8] especially when only shorter follow-up durations are available; and indeed more rapid structural loss has been reported to be predictive of changes in quality of life even after adjusting for the rate of functional loss.[9]

A key obstacle to greater clinical reliance on OCT as a tool to monitor progression is the presentation of the results to the clinician. Images of the scans typically take up only a small portion of the printout,[10, 11] and the automated delineation of those images used to generate objective structural measurements such as the retinal nerve fiber layer thickness (RNFLT) and Bruch's Membrane Opening Minimum Rim Width (MRW) is prone to error.[12–14] Analysis tools for progression are not yet as well developed for OCT[15] as for perimetry.[16] The clinician thus relies more on a subjective evaluation of the few quantifications of the image that are presented on the instrument's software and/or on the printouts. In the case of the Spectralis OCT (Heidelberg Engineering GmbH, Heidelberg, Germany) these consist of RNFLT and MRW, globally and in each of six predefined sectors.

In this study, we wished to examine how the rates of change of these quantifications relate to experienced clinicians' subjective assessment of the rate of structural progression. We conducted a survey of 11 international glaucoma experts, and derived an average score for the rate of structural progression in an individual eye based on its series of OCT printouts. This allows us to identify the quantifications of the rate of change that best correspond with this average score. This information can then be used to improve presentation of the data for clinical use; and to inform less experienced clinicians and researchers about the best ways to use OCT data. In so doing, we also obtain an invaluable dataset that can be used to derive and validate alternative quantifications of the rate of change, for example via machine learning techniques. Finally, it reveals features of OCT data that are currently thought to be the most relevant for assessing progression. Improving assessment of the rate of structural change will improve the ability of clinicians to integrate this with other relevant information when making clinical decisions.

Methods

Data for this study were taken from the ongoing longitudinal Portland Progression Project.[17–19] Study participants are enrolled if they have, or are at elevated risk of developing, primary open angle glaucoma, as determined by their clinician. Eyes that have visual field loss due to any non-glaucomatous pathology are excluded. Otherwise, no formal criteria are imposed based on intraocular pressure or the presence of a visual field or structural defect, so that the dataset encompasses glaucoma suspect eyes as well as those with clear glaucomatous damage. In particular, both eyes are tested even if glaucoma is unilateral. Participants undergo an array of structural and functional testing once every six months including OCT scans and automated perimetry. Participants gave informed consent before testing at every visit. All research and data sharing protocols were approved by the Legacy Health Institutional Review Board and adhered to the tenets of the Declaration of Helsinki.

The structural testing relevant to this study was conducted on the Heidelberg Spectralis OCT2. At each visit, a 6° radius circle scan is conducted centered on the centroid of Bruch's Membrane Opening (BMO), from which RNFLT is calculated. Twenty-four radial scans are also conducted, centered on the BMO centroid and aligned relative to the axis from that centroid to the fovea, from which MRW is calculated. Obvious delineation errors affecting RNFLT or MRW measurements were corrected on the instrument by the experienced technician prior to export of the data.[12, 13] The data were anonymized, then the instrument's software was used to generate the one page "Minimum Rim Width & RNFL Analysis Single Exam Report" for each visit date. This shows radial images of the neuroretinal rim at 12 locations around the rim, with MRW marked; a plot of the MRW profile around the rim, with normative limits shown; a plot of the 6° radius RNFLT profile around the nerve head, with normative limits shown; and the values of MRW and RNFLT averaged globally and averaged within each of six sectors[20] together with the corresponding percentile from the age-corrected normative database.[21] The instrument's software was also used to generate the "RNFL Change Report, All Follow-Ups", which shows for each test date images and plots of the 6° radius RNFLT profile around the optic nerve head with normative limits shown, and displays the same RNFLT values averaged globally and within each of six sectors.

For this study, structural testing data from the five most recent visits were used for each series. Eyes that underwent surgery other than uncomplicated cataract surgery during this period, were excluded. Although eyes were not excluded if they underwent cataract surgery, the date of any such surgery was noted on the printout.

A total of 120 longitudinal series were assembled in this way, from 100 eyes, with 20 eyes being repeated. The full set of printouts (the five single exam reports plus the RNFL change report) were then sent to a group of 11 experienced glaucoma experts for evaluation. Clinicians were chosen such that no more than two studied or practice at any one institution, to ensure a representative diversity of opinion. Each clinician independently gave a score for the rate of structural progression for each eye, with the following guidance: 1=Improvement, 2=Stable, 3=Very Slow, 4=Slow, 5=Medium. 6=Rapid, 7=Very Rapid. Scores from the 11 clinicians were averaged, to provide a reliable and effectively continuous metric of the rate of progression reflecting current best clinical practice. The validity of averaging scores and implicitly assuming that they constitute a continuous linear scale was assessed using an ordinal regression model (see Supplemental Figure 1 for details, available at http://www.aaojournal.org).[22]

Weighted generalized estimating equation (GEE) models were used to assess potential predictors of the mean clinician score. Predictors that were considered included the rates of change of RNFLT or MRW (based on the values presented on the clinical printouts), either globally or in the most rapidly changing sector; in each case using either raw values in μm/y or after dividing by the average value in the series to obtain rates in %/y. For the regression analysis, series were weighted proportional to the reciprocal of the within-series variance, estimated as the sum of the intra-eye within-clinician variance for the 20 repeated series plus the intra-eye between-clinician variance for that particular series. GEE regression was used in order to adjust for the presence of 20 repeated eyes. Goodness of fit of the models was assessed using the pseudo-R^2 for GEE regression developed by Zheng,[23] with 95% confidence intervals for the pseudo-R^2 values derived by bootstrapping with 1000 random resamplings of the dataset. Analyses were performed using R version 4.0.0,[24] with the geepack library.

Results

The 100 eyes came from 51 individuals, of whom 26 (51%) were female. Self-reported ethnicity was white for 47 individuals; 1 was Asian, and 3 individuals had more than one ethnicity recorded. The mean age was 72 years (range 49–88). Table 1 shows the characteristics of the eyes in the dataset, together with summaries of each metric of the rate of change derived from the instrument's quantifications of RNFLT and MRW. Values shown are for the 100 unique series, i.e. the 20 repeated series are not counted twice in these summaries. The mean image quality for the 6° radius circle scan was 29.9; one scan had a quality score of 19, and 24 scans had quality scores below 25.[25]

Among the 20 series that appeared twice in the dataset, the scores assigned by individual clinicians were identical 66% of the time, and within one point for 95% of these series. The intra-class correlation coefficient (ICC) for the mean score was 0.89 (95% confidence

interval 0.85 to 0.91), suggesting that this provides a highly reliable metric of clinicians' judgment of the rate of structural progression.

Among all 120 series, the average clinician score was 2.6 (on the 1–7 scale). Mean scores ranged from 1.5 to 4.8. Although there were 16 scores of 6 (rapid) or 7 (very rapid) given by individual clinicians, no series received these high scores from more than three clinicians. The intra-eye between-clinician standard deviations for the 120 series averaged 0.69 points.

Figure 1 shows scatterplots of the mean of the 11 clinician scores plotted against the rate of RNFLT change, either globally (top row) or in the most rapidly thinning sector (bottom row); when change is expressed in μm/y (left column) or as %/y (right column). Figure 2 shows the equivalent scatterplots for MRW. Table 2 shows the goodness-of-fit of univariate weighted GEE models (as described in the Methods section) for each of the quantitative structural predictors, using the pseudo-R^2 for GEE models. The average clinician score was much more closely correlated with the rate of thinning of RNFLT than MRW, as seen by the fact that the pseudo-R^2 values for RNFLT are above the 95% confidence intervals for the MRW pseudo-R^2 values. The best predictor of the mean clinician score was the rate of change of RNFLT in percent per year in the most rapidly thinning sector. This prediction equation was Mean Score = 1.93 − 0.20 × Sectoral Rate, where Sectoral Rate is for the worst sector in %/y; the 95% confidence interval for the regression coefficient was (−0.25, −0.15). The predicted average clinician score was ≥3, i.e. "Very Slow" progression or worse, when the most rapidly change sector had a rate worse than −5.2%/y.

Figure 3 shows ROC curves for whether each of the quantitative structural predictors accurately predicted whether the average clinician score would be ≥3. The areas under the full ROC curve were 0.87 for the global rate of RNFLT in μm/y; 0.89 for global rate in %/y; 0.90 for the worst sector rate in μm/y; and 0.92 for the worst sector rate in %/y. The partial areas under the curves, restricted to the specificities ≥80% that are most clinically relevant, were 0.11, 0.12, 0.14 and 0.15 respectively. Areas under the curve were consistently lower when using MRW instead of RNFLT.

As seen in Table 1 and the top left plot in Figure 1, one eye had an apparent thickening of the RNFLT of 15μm/y. This was caused by a schisis occurring in the superior nasal quadrant that crossed the 6° radius circle scan.[26] A further eye had an apparent thickening of 9.7%/y for RNFLT, as seen in the top right plot in Figure 1 (this was one of the eyes whose series appeared twice in the dataset hence appears twice on the plot); this was caused by fluctuation in the layer delineations in the superior temporal quadrant in an eye that had an absolute localized scotoma at the corresponding inferior temporal region of the visual field. The right-hand column of Table 2 shows the fits from the same univariate models as before but omitting these outliers; this did not materially change the results. The analysis was also repeated using the log of the mean clinician score instead of the native scaled score, in order to reduce the effect of the possible non-linearity in the relation seen in Figure 1, but this did not alter any comparisons between the fits using the different quantitative structural predictors (full results not shown).

Scores assigned by individual clinicians to each series were analyzed in the same way. The rate of change of RNFLT in the worst sector in %/y was the best predictor of score for 8 of the 11 clinicians. For two of the remaining clinicians, the rate for RNFLT in the worst sector in μm/y gave a slightly higher goodness of fit (by margins of 0.04 and 0.01); for one clinician the best predictor was the rate of change in global MRW.

The primary analysis relies on averaging scores among the 11 clinicians, implicitly assuming that scores exist on an evenly-spaced continuous linear scale. This assumption provides greater statistical power than the alternative of assuming scores exist on an ordered discrete scale, by reducing the number of free variables; and facilitates easier interpretation of the results. To validate this assumption, an ordinal regression model was used; see Supplemental Figure 1 for details (available at http://www.aaojournal.org). Although the assumption appears to be acceptable, there was some evidence suggesting that the difference between scores of 1 (improvement) and 2 (stable) was greater than the difference between other consecutive scores. This suggests that clinicians were reluctant to assign a score of 1, and required stronger evidence before concluding that an eye appeared to be improving. Thus, the primary analysis was repeated excluding 14 series for which at least one clinician gave a score of 1. This did not materially alter the conclusions. Reanalyzing the remaining 106 series, the goodness of fit (as in Table 2) was 0.645 for the rate of change of RNFLT in the worst sector in %/y; 0.557 for the worst sector in μm/y; 0.546 for the global rate in %/y; and 0.512 for the global rate in μm/y; in each case these are within the 95% confidence intervals for the corresponding values using the full dataset.

Figure 4 shows the goodness of fit (using the pseudo-R^2 as before) using each quantitative structural metric of rate to predict the mean clinician score, in four subsets of the data: 16 series whose average RNFLT within the series was <60μm; 48 series with average RNFLT ≥60μm but <80μm; 40 series with average RNFLT ≥80μm but <100μm; and 13 series with average RNFLT ≥100μm (excluding the same three outlier series as before). The rate of thinning of RNFLT in the most rapidly changing sector was the best predictor of average clinician score across the range of disease severities.

The “Worst Sector Rate” in the above results did not occur in the same sector for every eye. Table 3 shows how often each of the six instrument-defined sectors was the one that had the most rapid change in %/y, and so was the sector whose rate of thinning was used in the analysis. The location with the most rapid thinning was relatively evenly spread around the optic nerve head, although it occurred slightly more frequently in the superior temporal sector for both RNFLT and MRW. Notably, the rate in %/y of the most rapidly thinning sector of RNFLT was a better predictor of mean clinician score ($R^2 = 0.667$) than any individual sector or pair of sectors (the highest R^2 achieved was 0.523 when using the temporal and inferior nasal sectors as predictors).

Some OCT instruments from other manufacturers report RNFLT in just two sectors, “superior” and “inferior”. When combining the Spectralis’ sectors in this way, the rate in the superior sector predicted mean clinician score with $R^2 = 0.270$; the rate in the inferior sector gave $R^2 = 0.459$; and using the worse of these two sectors in an eye gave $R^2 = 0.514$, significantly lower than using the worst rate among six smaller sectors.

Discussion

In this study, we found that the subjective assessment of the rate of structural progression by glaucoma specialists using OCT scans was more closely related to the rate of localized change in the fastest deteriorating sector for that particular eye than to the global rate of change. Furthermore, the clinicians' score was more highly correlated with the rate in %/y than in μm/y; that is, the same amount of tissue loss in μm/y was considered to constitute more rapid disease progression if the eye already had a thinner nerve fiber layer, and/or if it occurred in a region within the eye that already had a thinner nerve fiber layer. This rate was the best predictor of the average score among the 11 clinicians, and also for 8 of the 11 individual clinicians; and performed consistently well across disease severities. A rate worse than −5.2%/y corresponded with an average score of "very slow progression" or worse. Finally, we found that the clinicians' scores were more closely related to the rate of reduction in RNFLT than in MRW.

Given that RNFLT and MRW comprise both neural and non-neural tissue, in proportions that may vary between eyes and over the course of the disease,[27, 28] it is not possible to determine whether the clinicians are correct in their judgment of the true rate of progression; such studies will only become possible when we have the ability to perform in vivo full-retina ganglion cell counting, which remains some years away. In the absence of such validation, the average judgment of experts remains a reasonable "clinical standard" assessment of the rate of progression and its clinical importance. This kind of "collective intelligence" approach is more repeatable than scores from any single clinician, since averaging multiple scores reduces variability.[29, 30] Indeed, the worst sectoral proportional rate of RNFLT change predicted the mean clinician score with a pseudo-R^2 of 0.657, as seen in Table 2; whereas the highest value seen for any individual clinician was 0.637. Using the average of multiple clinicians also increases the generalizability of our results, especially since we ensured that no more than two of the clinicians were from the same institution. Since the score is not tied to any individual metric of the rate of progression, it can be used in subsequent studies to develop and validate alternative quantifications of the rate, for example via machine learning approaches.

Anecdotally, many clinicians rely more on functional than structural information when assessing rate of disease progression in eyes with glaucoma. There are likely to be several reasons for this, but one is that the presentation of data to the clinician from OCT scans is not yet as well developed as it is for automated perimetry. In that regard, it is notable that sectoral rates of change are generally not actually presented on the current Spectralis printouts that were given to the clinicians; nor are proportional rates (i.e. rates expressed in %/y instead of μm/y). The sectoral rates (and associated p-values) are available through the instrument's software in some countries, although these are not yet federally approved and available in the USA. It is therefore even more impressive that this rate metric was the one that best correlated with clinicians' judgment. Our results suggest new metrics that may prove to be a worthwhile addition to standard clinical printouts.

Notably, sectoral rates of change are also not presented on the standard printouts from other manufacturers' OCT devices. The partial exception is the "superior" and "inferior" rates

that are reported by the Topcon Maestro (Topcon Corporation, Tokyo, Japan) and Cirrus (Carl Zeiss Meditec, Dublin, CA, USA) devices. However, rates in these larger sectors did not correlate as well with the clinicians' judgment, suggesting that they may be too wide to detect focal changes that are known to be characteristic of glaucoma. When using larger sectors or global analyses, variability in stable healthy regions can diminish the signal-to-noise ratio that would be achieved in more localized regions of interest.

It is also notable that it was the sector with the worst rate in %/y that was found to be the strongest predictor, rather than the sector with the worst rate in μm/y. RNFLT is usually greatest superiorly and inferiorly, and smallest nasally and temporally. Using the proportional rate in %/y therefore places greater emphasis on nasal and temporal thinning, as seen by the relatively uniform spread of sectors in Table 3. This can be helpful in particular for the temporal RNFLT, since much of this sector corresponds with the macular region that is often affected early in the disease process and may be most important for activities of daily living.[11] It should be noted however that the pseudo-R^2 value using the rate in %/y was within the 95% confidence interval for the pseudo-R^2 using the rate in μm/y (Table 2), indicating that this difference was not statistically significant; and that a difference was only observed in one of the four subsets of the data shown in Figure 4 (the subset containing 40 series with average RNFLT 80–100μm).

Rates of change of RNFLT consistently predicted clinician scores better than the equivalent rates of change for MRW. The information provided to the clinicians included the "Minimum Rim Width & RNFL Analysis Single Exam Report" for each of the five scan dates, on which around half of the available space is devoted to MRW and radial images relevant to MRW; while only around a quarter of the available space is devoted to RNFLT and the peripapillary circle scan image on which it is based. Clinicians were also provided with the "RNFL Change Report" printout, which shows the peripapillary circle scans from all five test dates. At present, no equivalent report is available for MRW. Thus one possible reason for the observed reliance on RNFLT more than MRW may be the way the data is presented, with clinicians using the change report printout (that only shows RNFLT) more than the single exam reports (which show both but emphasize MRW more). Another possible reason is that RNFLT has been available for longer, so clinicians are more familiar with its use. It should be noted though that RNFLT is less variable than MRW,[31] with resultant better diagnostic accuracy,[32] possibly due to the difficulty of automatically accurately detecting BMO. This variability will reduce correlations between the rate of change of MRW and the clinician scores. At the same time, the higher variability of MRW provides a sound logic for why clinicians *should* rely more on RNFLT at the current time, at least until the variability of MRW is reduced by improvements in the automated detection of BMO.

The OCT scans for this study were taken from a well-controlled longitudinal study (the Portland Progression Project), acquired by a small number of dedicated technicians who had sufficient time to obtain well-focused, well-centered images, repeating scans if necessary. As such, the scans are of higher quality than might be typical in a clinical situation. To best replicate clinical practice, image quality score was available on the printouts but was not a formal exclusion criterion. Even so, only one scan had an image quality score below 20; the

average quality score was 29.9. In a more typical clinical situation with less clear images, it would be natural for the clinician to rely even more heavily on the instrument software's quantifications of the images rather than visually examining the images themselves. Thus, it is possible that we are underestimating the correlation between clinician judgment and quantitative indices. However, lower quality images would also be expected to have higher test-retest variability.

A methodological caveat is that quantitative rates of change were compared based on their ability to predict average clinician score using a weighted GEE regression model. Such models are fit using quasi-likelihood estimation, and so a formal statistical test based on likelihood ratios comparing non-nested models that use different predictors is not appropriate. Models were compared using Zheng's pseudo-R^2 for GEE models,[23] but it is not possible to generate a p-value to determine whether one model is significantly better than another. Instead, 95% confidence intervals were generated by bootstrapping. This confirmed that the pseudo-R^2 for the worst sector rate of RNFLT was outside the 95% interval of the values for global RNFLT, global MRW, or sectoral MRW, suggesting a significant difference in each case with $p<5\%$. It should also be noted that our conclusions are supported by visual inspection of the plots in Figures 1 and 2.

A final caveat is that series presented to the clinicians were only 5 scans, covering a period of around two years. The number of rapidly progressing eyes was small, consistent with the distribution of rates that would be expected in a typical clinical population.[33] The highest average score of any series was only 4.8; only one eye showed more than 15μm of RNFLT thinning in that period, and clinicians may be reluctant to say that an eye is experiencing very rapid progression based on such a short series since apparent changes could just be due to variability. With longer series we may also have seen greater agreement between the clinicians. The short length of the series also means that the estimated rates themselves have large 95% confidence intervals. However, the challenging nature of assessing the rate of change in short series was part of the motivation for this study; this is a common situation in the clinic and improved tools to aid the clinician in this scenario are needed. It is possible that the relative utility of different quantifications may differ in very rapidly progressing eyes, or in eyes with more severe glaucoma.

In conclusion, we found that glaucoma specialists' assessment of the rate of structural glaucomatous progression is closely correlated with the rate of thinning in %/y of the most rapidly thinning sector in that eye of the peripapillary retinal nerve fiber layer. This quantitative metric of rate of loss performed better than the rate of change of the global average RNFLT; the rate in μm/y; or the rate of change in MRW. This best predictor had a pseudo-R^2 of 0.657 when predicting the average score of 11 clinicians, which is strong enough to be useful yet still leaves room for improvement, suggesting a possible role for using machine learning approaches to derive new metrics of the rate of change. In the meantime, this metric of rate of change can be calculated now from information already available to the clinician for some OCT devices. We would encourage instrument manufacturers to make localized quantifications more readily available both on the devices and in clinical printouts.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Financial Support:

This work was supported by the National Institutes of Health (grant numbers NEI R01 EY020922 and R01 EY031686 to author SKG); and the Good Samaritan Foundation. The funding organizations had no role in the design or conduct of this research.

Conflict of Interest:

SKG: Heidelberg Engineering (Research Support)

CGDM: Heidelberg Engineering, Carl Zeiss Meditec, Topcon (Research Support); Novartis, Reichart, Galimedix, Belite, Ora Clinical, Thea Pharma, Allergan (Consultant)

KNM: Heidelberg Engineering (Research Support)

JSM: Aerie, Allergan, Diopsys, Equinox, Glaukos, Haag Streit, Nicox, Olleyes, Santen (Research Support); Aerie, Allergan, Avisi, Glaukos, Olleyes (Consultant)

DLB: Carl Zeiss Meditec (Patent); Carl Zeiss Meditec (Consultant) SLM: Allergan (Research Support); Thea, Nicox (Consultant) Other authors have no conflicts of interest relevant to this work.

References

1. Freeman EE, Munoz B, West SK, Jampel HD, Friedman DS. Glaucoma and quality of life: the Salisbury Eye Evaluation. Ophthalmology 2008;115(2):233–8. [PubMed: 17655930]
2. McKean-Cowdin R, Wang Y, Wu J, Azen SP, Varma R, Los Angeles Latino Eye Study G. Impact of visual field loss on health-related quality of life in glaucoma: the Los Angeles Latino Eye Study. Ophthalmology 2008;115(6):941–8 e1. [PubMed: 17997485]
3. Medeiros FA, Gracitelli CP, Boer ER, Weinreb RN, Zangwill LM, Rosen PN. Longitudinal changes in quality of life and rates of progressive visual field loss in glaucoma patients. Ophthalmology 2015;122(2):293–301. [PubMed: 25444345]
4. Alqudah A, Mansberger SL, Gardiner SK, Demirel S. Vision-related Quality of Life in Glaucoma Suspect or Early Glaucoma Patients. J Glaucoma 2016;25(8):629–33. [PubMed: 27483331]
5. Medeiros FA, Tatham AJ. Structure versus Function in Glaucoma: The Debate That Doesn't Need to Be. Ophthalmology 2016;123(6):1170–2. [PubMed: 27210598]
6. Gardiner SK, Fortune B, Demirel S. Signal-to-Noise Ratios for Structural and Functional Tests in Glaucoma. Transl Vis Sci Technol 2013;2(6):3.
7. Zhang X, Dastiridou A, Francis BA, et al. Comparison of Glaucoma Progression Detection by Optical Coherence Tomography and Visual Field. Am J Ophthalmol 2017;184:63–74. [PubMed: 28964806]
8. Abe RY, Diniz-Filho A, Zangwill LM, et al. The Relative Odds of Progressing by Structural and Functional Tests in Glaucoma. Invest Ophthalmol Vis Sci 2016;57(9):OCT421–OCT8. [PubMed: 27409501]
9. Gracitelli CP, Abe RY, Tatham AJ, et al. Association between progressive retinal nerve fiber layer loss and longitudinal change in quality of life in glaucoma. JAMA Ophthalmol 2015;133(4):384–90. [PubMed: 25569808]
10. Hood DC, De Cuir N, Blumberg DM, et al. A Single Wide-Field OCT Protocol Can Provide Compelling Information for the Diagnosis of Early Glaucoma. Transl Vis Sci Tech 2016;5(6):4.
11. Hood DC, De Moraes CG. Challenges to the Common Clinical Paradigm for Diagnosis of Glaucomatous Damage With OCT and Visual Fields. Investigative ophthalmology & visual science 2018;59(2):788–91. [PubMed: 29392325]

12. Menda S, Fortune B, Mansberger SL, Gardiner SK, Demirel S. The effect of manually refining automated image segmentations on retinal nerve fiber layer thickness measurements from optical coherence tomography (OCT) scans. Invest Ophthalmol Vis Sci 2015;56(7):4551.
13. Mansberger SL, Menda SA, Fortune BA, Gardiner SK, Demirel S. Automated Segmentation Errors When Using Optical Coherence Tomography to Measure Retinal Nerve Fiber Layer Thickness in Glaucoma. Am J Ophthalmol 2017;174:1–8. [PubMed: 27818206]
14. Nagarkatti-Gude N, Gardiner SK, Fortune B, Demirel S, Mansberger SL. Optical Coherence Tomography Segmentation Errors of the Retinal Nerve Fiber Layer Persist Over Time. J Glaucoma 2019;28(5):368–74. [PubMed: 30855415]
15. Dong ZM, Wollstein G, Schuman JS. Clinical Utility of Optical Coherence Tomography in Glaucoma. Invest Ophthalmol Vis Sci 2016;57(9):OCT556–OCT67. [PubMed: 27537415]
16. Heijl A, Lindgren G, Lindgren A, et al. Extended empirical statistical package for evaluation of single and multiple fields in glaucoma: Statpac 2. In: Mills R, Heijl A, editors. Perimetry Update 1990/1991. Amsterdam: Kugler & Ghedini, 1991:303–15.
17. Gardiner SK, Johnson CA, Demirel S. Factors predicting the rate of functional progression in early and suspected glaucoma. Invest Ophthalmol Vis Sci 2012;53(7):3598–604. [PubMed: 22570353]
18. Gardiner SK, Mansberger SL, Demirel S. Detection of Functional Change Using Cluster Trend Analysis in Glaucoma. Invest Ophthalmol Vis Sci 2017;58(6):Bio180–Bio90. [PubMed: 28715580]
19. Gardiner SK, Mansberger SL, Fortune B. Time Lag Between Functional Change and Loss of Retinal Nerve Fiber Layer in Glaucoma. Invest Ophthalmol Vis Sci 2020;61(13):5-.
20. Garway-Heath DF, Poinoosawmy D, Fitzke FW, Hitchings RA. Mapping the visual field to the optic disc in normal tension glaucoma eyes. Ophthalmology 2000;107(10):1809–15. [PubMed: 11013178]
21. Chauhan BC, Danthurebandara VM, Sharpe GP, et al. Bruch's Membrane Opening Minimum Rim Width and Retinal Nerve Fiber Layer Thickness in a Normal White Population: A Multicenter Study. Ophthalmology 2015;122(9):1786–94. [PubMed: 26198806]
22. Bürkner P-C, Vuorre M Ordinal Regression Models in Psychology: A Tutorial. Advances in Methods and Practices in Psychological Science 2019;2(1):77–101.
23. Zheng B Summarizing the goodness of fit of generalized linear models for longitudinal data. Statistics in medicine 2000;19(10):1265–75. [PubMed: 10814976]
24. R Development Core Team. R: A language and environment for statistical computing, 4.0.0 ed. Vienna, Austria: R Foundation for Statistical Computing, 2020.
25. Balasubramanian M, Bowd C, Vizzeri G, Weinreb RN, Zangwill LM. Effect of image quality on tissue thickness measurements obtained with spectral domain-optical coherence tomography. Optics express 2009;17(5):4019–36. [PubMed: 19259243]
26. Fortune B, Ma KN, Gardiner SK, Demirel S, Mansberger SL. Peripapillary Retinoschisis in Glaucoma: Association With Progression and OCT Signs of Müller Cell Involvement. Invest Ophthalmol Vis Sci 2018;59(7):2818–27. [PubMed: 29860466]
27. Harwerth RS, Wheat JL, Fredette MJ, Anderson DR. Linking structure and function in glaucoma. Prog Retin Eye Res 2010;29(4):249–71. [PubMed: 20226873]
28. Patel NB, Sullivan-Mee M, Harwerth RS. The relationship between retinal nerve fiber layer thickness and optic nerve head neuroretinal rim tissue in glaucoma. Invest Ophthalmol Vis Sci 2014;55(10):6802–16. [PubMed: 25249610]
29. Barnett ML, Boddupalli D, Nundy S, Bates DW. Comparative Accuracy of Diagnosis by Collective Intelligence of Multiple Physicians vs Individual Physicians. JAMA Network Open 2019;2(3):e190096. [PubMed: 30821822]
30. Wolf M, Krause J, Carney PA, Bogart A, Kurvers RH. Collective intelligence meets medical decision-making: the collective outperforms the best radiologist. PLoS One 2015;10(8):e0134269. [PubMed: 26267331]
31. Gardiner SK, Boey PY, Yang H, Fortune B, Burgoyne CF, Demirel S. Structural Measurements for Monitoring Change in Glaucoma: Comparing Retinal Nerve Fiber Layer Thickness With Minimum Rim Width and Area. Invest Ophthalmol Vis Sci 2015;56(11):688–691.

32. Stagg BC, Medeiros FA. A Comparison of OCT Parameters in Identifying Glaucoma Damage in Eyes Suspected of Having Glaucoma. Ophthalmology Glaucoma 2020;3(2):90–6. [PubMed: 32632407]

33. Chauhan BC, Malik R, Shuba LM, Rafuse PE, Nicolela MT, Artes PH. Rates of glaucomatous visual field change in a large clinical population. Invest Ophthalmol Vis Sci 2014;55(7):4135–43. [PubMed: 24917147]

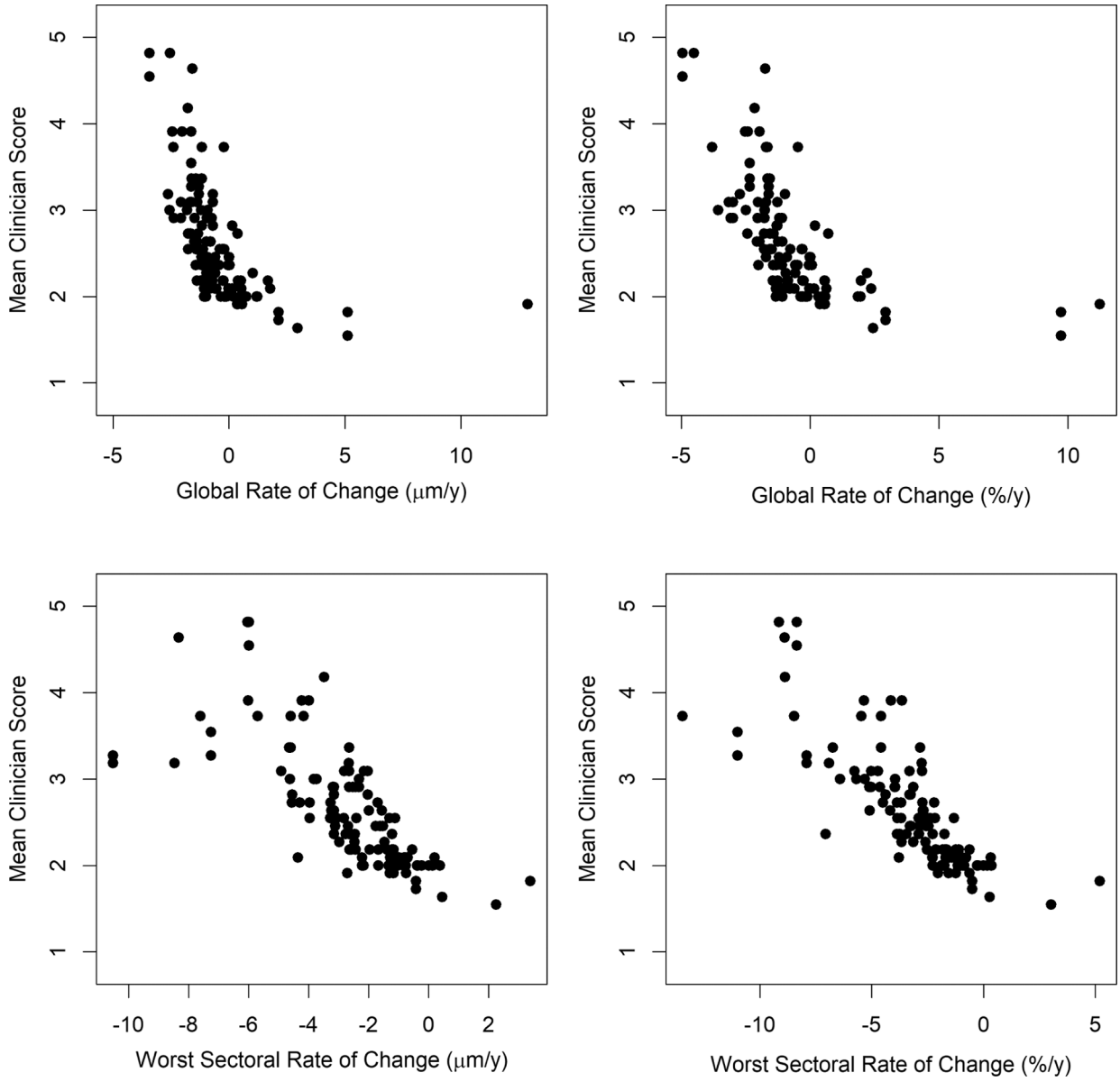

Figure 1:

Clinicians' subjective assessment of the rate of structural progression, on a scale from 1 (improvement) to 7 (very rapid progression), averaged between 11 glaucoma specialists, plotted against the rate of change of Retinal Nerve Fiber Layer Thickness (RNFLT). Series consisted of five biannual scans. Top row: global average RNFLT. Bottom row: average RNFLT within the most rapidly thinning of the six sectors presented on the instrument's software. Left column: rate expressed as microns per year. Right column: rate expressed as percent per year of the mean value within the series.

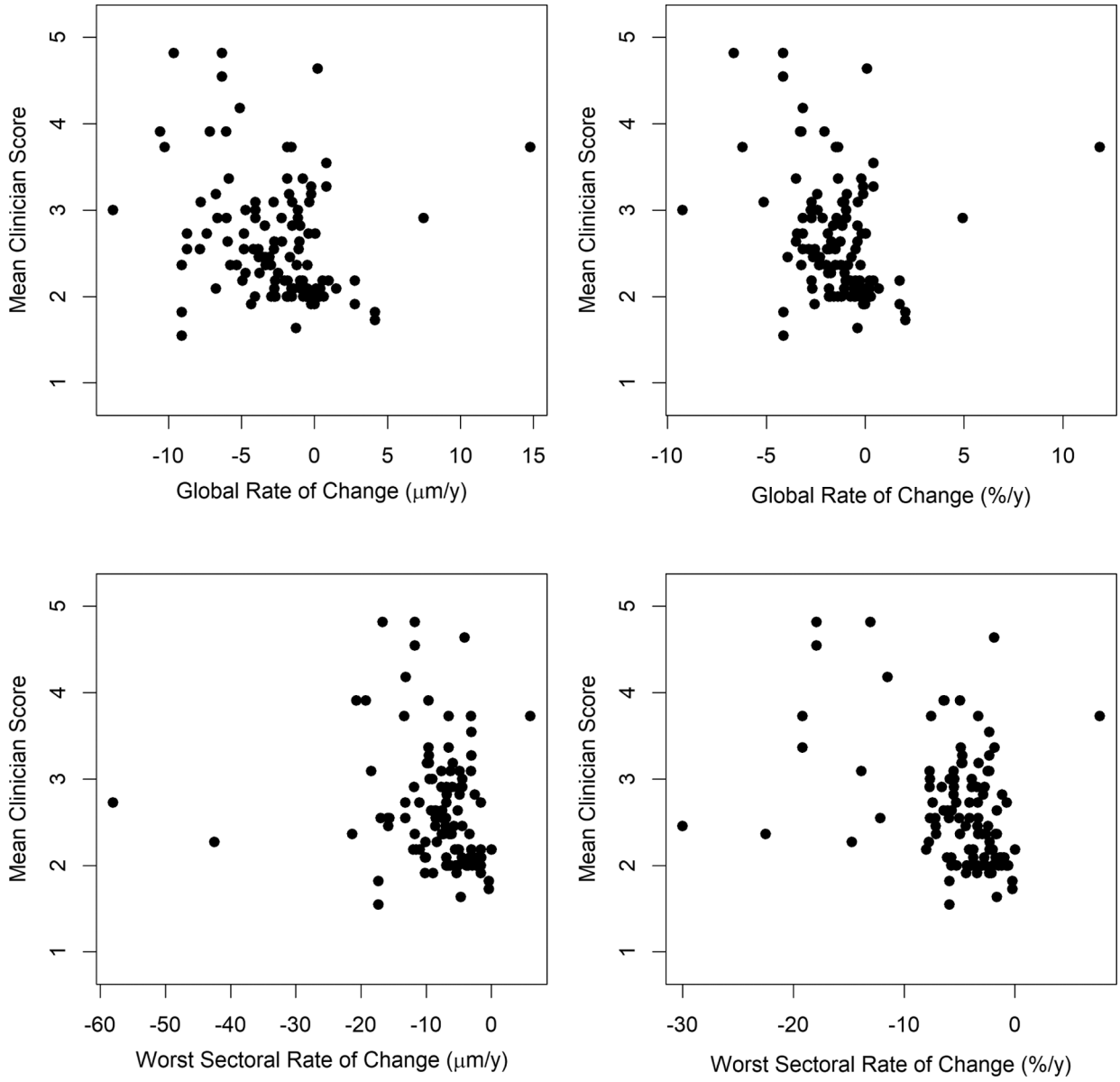

Figure 2:

Clinicians' subjective assessment of the rate of structural progression, on a scale from 1 (improvement) to 7 (very rapid progression), averaged between 11 glaucoma specialists, plotted against the rate of change of Bruch's Membrane Opening Minimum Rim Width (MRW). Series consisted of five biannual scans. Top row: global average MRW. Bottom row: average MRW within the most rapidly thinning of the six sectors presented on the instrument's software. Left column: rate expressed as microns per year. Right column: rate expressed as percent per year of the mean value within the series.

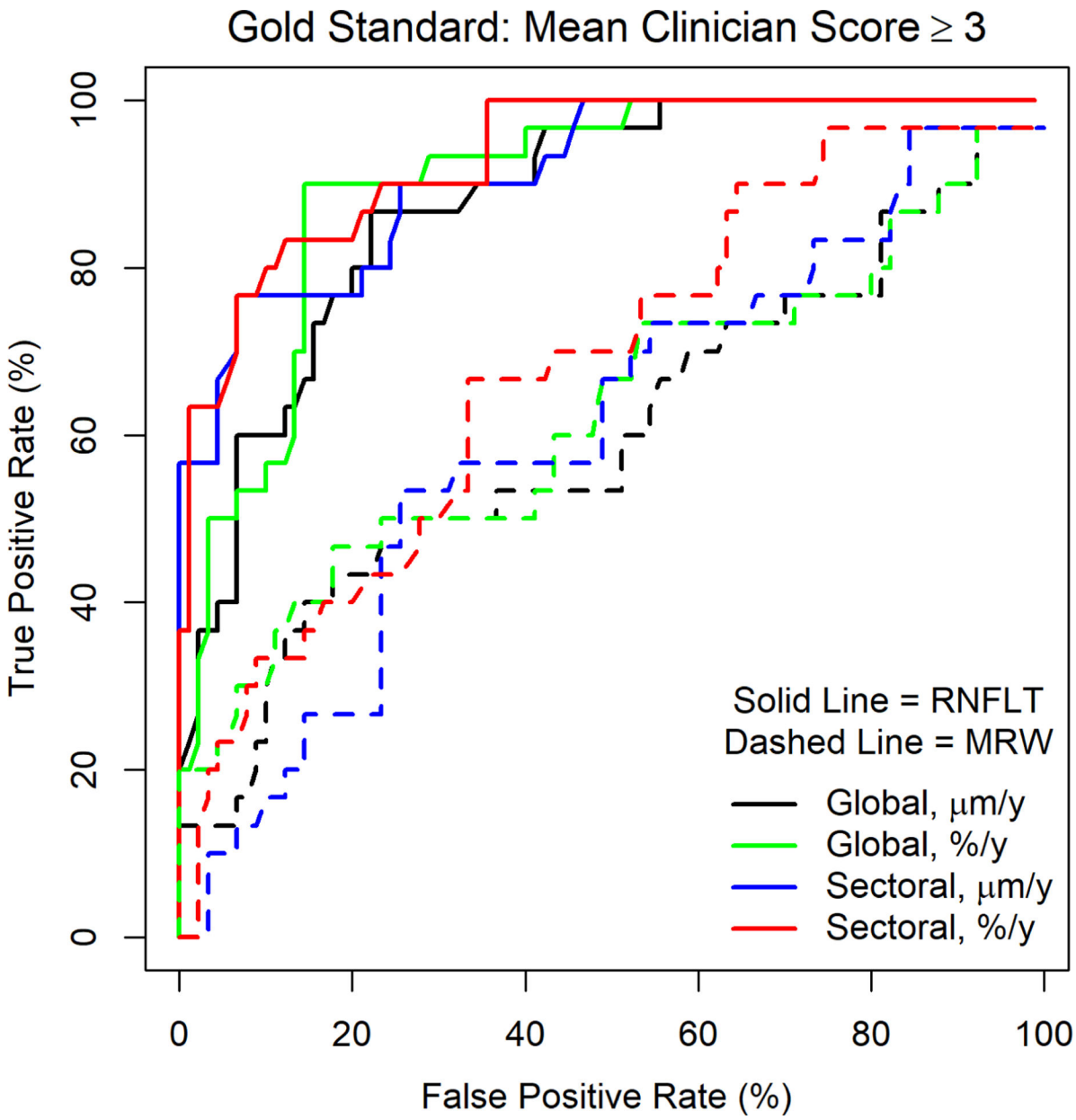

Figure 3:

Receiver Operating Characteristic (ROC) Curves, showing the sensitivity and specificity of each quantitative measure of the rate of change when predicting whether the average subjective assessment of expert clinicians would be a score of 3 (“Very Slow Progression”) or worse. Measures considered were the global average rate, and the rate in the most rapidly thinning of six sectors, for Retinal Nerve Fiber Layer Thickness (RNFLT) and Bruch’s Membrane Opening Minimum Rim Width (MRW).

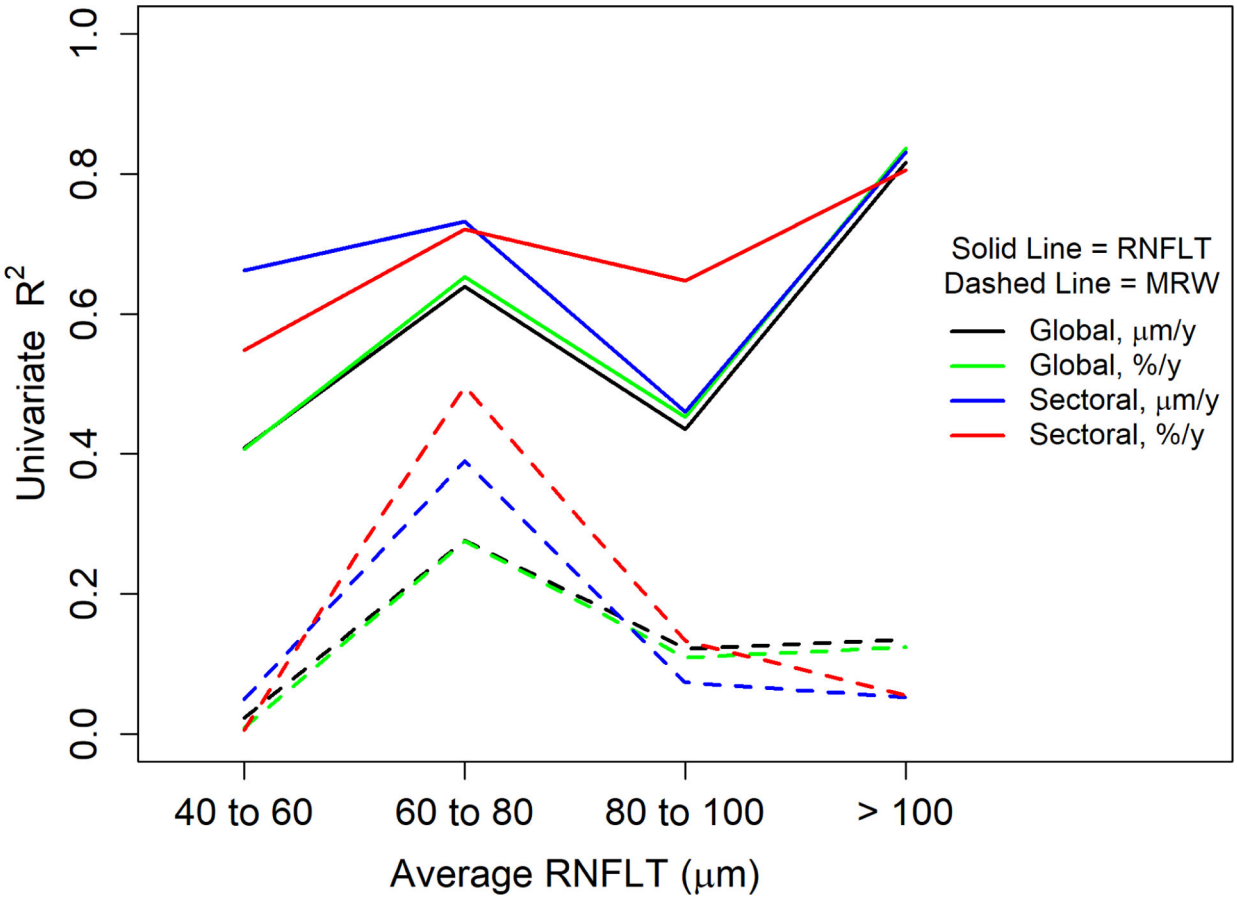

Figure 4:

Goodness of fit for univariate models predicting clinicians' subjective assessment of the rate of structural progression from various quantitative structural metrics of the rate of change, in subsets of the data defined by the mean Retinal Nerve Fiber Layer Thickness (RNFLT) within the series. Clinicians' scores were on a scale from 1 (improvement) to 7 (very rapid progression), averaged between 11 glaucoma specialists. Goodness of fit was measured using the pseudo-R^2 for weighted generalized estimating equations, which can be thought of as representing the proportion of variance explained by the model. Solid lines are for univariate models based on RNFLT; dashed lines are for univariate models based on Minimum Rim Width (MRW). In each case, models use the rate of change of the global average in μm/y (black) or %/y (green); of the rate of change of the most rapidly thinning sector in μm/y (blue) or %/y (red).

Table 1:

Characteristics of the dataset. For the 100 unique series of five fields, the table shows the average value and rate of change of Mean Deviation (MD) and Visual Field Index (VFI) from automated perimetry; the average value and rates of change (in microns per year and percent per year) for global average Retinal Nerve Fiber Layer Thickness (RNFLT) and Minimum Rim Width (MRW); and the equivalent rates of change for the sectors exhibiting the most rapid change in that eye during the series.

	Mean	Interquartile Range	Range
Functional Parameters:			
Average MD (dB)	−2.4	−3.1 to +0.5	−16.8 to +2.8
Average VFI (%)	91.8	91.5 to 99.1	51.0 to 100.0
Rate of MD (dB/y)	−0.12	−0.46 to +0.27	−2.43 to +1.72
Rate of VFI (%/y)	−0.48	−1.02 to +0.18	−6.68 to +4.29
RNFLT:			
Global Average (μm)	79.3	67.9 to 93.8	41.4 to 126.6
Global Rate (μm/y)	−0.50	−1.41 to −0.02	−3.43 to +12.88
Global Rate (%/y)	−0.70	−1.74 to −0.00	−4.95 to +11.24
Worst Sector Rate (μm/y)	−2.62	−3.27 to 1.22	−10.52 to +3.39
Worst Sector Rate (%/y)	−3.21	−4.41 to −1.55	−13.46 to +5.20
MRW:			
Global Average (μm)	214.4	159.6 to 257.1	52.0 to 404.6
Global Rate (μm/y)	−2.60	−4.43 to −0.55	−13.79 to +14.79
Global Rate (%/y)	−1.29	−2.35 to −0.24	−9.24 to +11.87
Worst Sector Rate (μm/y)	−7.96	−9.91 to −4.37	−58.04 to +5.98
Worst Sector Rate (%/y)	−4.69	−5.97 to −1.95	−30.00 to +7.66

Table 2:

Goodness of fit of univariate models predicting clinicians' subjective assessment of the rate of structural progression from various quantitative structural metrics of the rate of change, based on global measures or on the most rapidly changing sector for that particular eye. Clinicians' scores were on a scale from 1 (improvement) to 7 (very rapid progression), averaged between 11 glaucoma specialists. Goodness of fit was measured using the pseudo-R^2 for weighted generalized estimating equations, which can be thought of as representing the proportion of variance explained by the model. 95% confidence intervals for these pseudo-R^2 values from bootstrapping are shown in parentheses.

Predictor	Using All Series	Excluding Outliers
Retinal Nerve Fiber Layer Thickness:		
Global Rate (μm/y)	0.305 (0.210 −0.557)	0.514 (0.381 −0.580)
Global Rate (%/y)	0.372 (0.286 −0.549)	0.543 (0.331 −0.603)
Worst Sector Rate (μm/y)	0.582 (0.497 −0.683)	0.579 (0.489 −0.686)
Worst Sector Rate (%/y)	0.657 (0.557 −0.749)	0.667 (0.572 −0.757)
Minimum Rim Width:		
Global Rate (μm/y)	0.084 (0.017 −0.256)	0.123 (0.017 −0.273)
Global Rate (%/y)	0.084 (0.016 −0.296)	0.114 (0.015 −0.319)
Worst Sector Rate (μm/y)	0.047 (0.017 −0.165)	0.063 (0.016 −0.177)
Worst Sector Rate (%/y)	0.149 (0.042 −0.365)	0.160 (0.040 −0.356)

Table 3:

Frequencies with which each of the six sectors was the sector with the most rapid thinning of Retinal Nerve Fiber Layer Thickness (RNFLT) or Minimum Rim Width (MRW), when expressed as percent per year of the mean value for that eye across the series, for the 100 series of five scans in the study.

Sector	RNFLT	MRW
Temporal	13	15
Superior Temporal	22	30
Superior Nasal	15	11
Nasal	20	13
Inferior Nasal	16	12
Inferior Temporal	14	19