Title

Evaluation of Progressive Neuroretinal Rim Loss as a Surrogate End Point for Development of Visual Field Loss in Glaucoma

Journal

Ophthalmology, 121(1)

ISSN

0161-6420

Authors

Medeiros, Felipe A
Lisboa, Renato
Zangwill, Linda M
et al.

Publication Date

2014

DOI

10.1016/j.ophtha.2013.06.026

Peer reviewed

Published in final edited form as:
Ophthalmology. 2014 January ; 121(1): . doi:10.1016/j.ophtha.2013.06.026.

Evaluation of Progressive Neuroretinal Rim Loss as a Surrogate Endpoint for Development of Visual Field Loss in Glaucoma

Felipe A. Medeiros, M.D., Ph.D.[1], **Renato Lisboa, M.D.**[1], **Linda M. Zangwill, Ph.D.**[1], **Jeffrey M. Liebmann, M.D.**[2], **Christopher A. Girkin, M.D., MPH**[3], **Christopher Bowd, Ph.D.**[1], and **Robert N. Weinreb, M.D.**[1]

[1]Hamilton Glaucoma Center, Department of Ophthalmology, University of California, San Diego

[2]New York Eye and Ear Infirmary, New York, NY, USA

[3]Department of Ophthalmology, University of Alabama, Birmingham, AL, USA

Abstract

Purpose—To evaluate the validity of using progressive neuroretinal rim area loss as a surrogate endpoint for development of visual field loss in glaucoma.

Design—Prospective observational cohort study.

Participants—The study group included 492 eyes of 328 patients classified as suspected of having glaucoma at the baseline visit. These eyes had an average of 7.4 ± 2.8 confocal scanning laser ophthalmoscopy (CSLO) images during a mean follow-up time of 6.6 ± 1.6 years.

Methods—Rim area measurements were acquired with CSLO during follow-up. The visual field endpoint was considered as development of 3 consecutive abnormal visual fields on standard automated perimetry. Strong predictive ability and large proportion of treatment effect explained (PTE) are requisites for a suitable surrogate endpoint. A joint longitudinal survival model was used to evaluate the ability of rates of rim area loss in predicting visual field development, adjusting for confounding variables (baseline age, race and corneal thickness, and follow-up measurements of intraocular pressure [IOP] and pattern standard deviation). The PTE was calculated comparing the effect of IOP on the risk of development of visual field loss when incorporating rim area loss in the same model versus the effect of IOP in the model excluding rim area measurements.

Main Outcome Measures—Predictive strength measured by survival-adapted R^2 and PTE.

Results—Sixty-two of 492 (13%) eyes developed visual field loss during follow-up. The mean rate of rim area change in eyes that developed visual field loss was −0.011mm^2/year versus −0.003mm^2/year in those that did not (P <0.001). In the multivariable model, each 0.01mm^2/year faster rate of rim area loss was associated with a 2.94 higher risk of visual field loss (hazard ratio = 2.94; 95% confidence interval: 1.38 – 6.23; P = 0.005). R^2 values were 62% and 81% for univariable and multivariable models, respectively. The PTE was 65%.

Corresponding Author: Felipe A. Medeiros, M.D., Ph.D., Hamilton Glaucoma Center, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0946, fmedeiros@glaucoma.ucsd.edu.

Financial Disclosures: Research support from Carl-Zeiss Meditec (FAM, LMZ, CAG, JML, RNW). Research support from Heidelberg Engineering (FAM, LMZ, RNW, JML). Consultant to Carl-Zeiss Meditec, Inc. (RNW, JML). None: RL, CB.

Conclusion—Progressive rim area loss was highly predictive of development of visual field loss in glaucoma and explained a significant proportion of treatment effect on the clinically relevant outcome. These findings suggest that rim area measurements may be suitable surrogate endpoints in glaucoma clinical trials.

Glaucoma is characterized by progressive structural changes of the optic nerve and retinal nerve fiber layer that may lead to loss of visual function and decreased vision-related quality of life. The fundamental goal of glaucoma management is to prevent patients from developing visual impairment that is sufficient to produce disability in their daily lives and impair their health-related quality of life.[1] However, due to the slowly progressive course of glaucoma, direct observation of disability endpoints is generally impractical for clinical trials testing new treatments for the disease.

When the use of the true endpoint is difficult and increases the complexity and duration of trials, an attractive solution is to replace the true endpoint by a biomarker that can be measured earlier, more conveniently or more frequently. A biomarker can be defined as a characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes, or pharmacologic responses to a therapeutic intervention.[2] According to this definition, several measures could potentially qualify as biomarkers in glaucoma, such as intraocular pressure, genetic markers, biochemical blood measurements, visual field assessment using standard perimetry, or imaging measurements of the optic disc and retinal nerve fiber layer. However, although many biomarkers can be associated with a disease, only a few potentially qualify as surrogate endpoints. According to the US Food and Drug Administration (FDA), "a surrogate end point is a laboratory measurement or physical sign that is used in therapeutic trials as a substitute for a clinically meaningful end point that is a direct measure of how a patient feels, functions, or survives and is expected to predict the effect of the therapy."[2] Therefore, in order to qualify as a surrogate endpoint, a biomarker needs to demonstrate significant ability to predict the clinically relevant outcome as well as the effect of treatment on this outcome.[3]

Although intraocular pressure (IOP) has traditionally been used as a surrogate endpoint in clinical trials, it is an imperfect surrogate for the clinically relevant outcomes of the disease. Many patients can progress despite low IOP levels and others remain stable despite having IOP measurements consistently high.[4, 5] Further, IOP is not a suitable surrogate endpoint for clinical trials investigating certain treatment modalities for glaucoma, such as neuroprotective therapies. The use of visual fields as the sole endpoint in glaucoma trials is also potentially limited by the need for large samples, long-term follow-up and variability of results.[6] In the past two decades, strong evidence has accumulated with regard to the role of structural measurements of the optic disc topography and retinal nerve fiber layer (RNFL) for diagnosing and detecting glaucoma progression. There is now substantial evidence that many patients can develop structural changes before appearance of detectable change in functional measures.[7–10] The use of structural measurements as surrogate endpoints in glaucoma clinical trials could potentially have a number of advantages, including faster acquisition of a sufficient number of endpoints with potential reduction in sample size requirements, enabling shorter, more effective, and less expensive trials.

The presence of a significant correlation between a biomarker and the clinically relevant endpoint is not sufficient to characterize a good surrogate.[11] For a surrogate endpoint to be useful in practice, it should be able to predict the effect of treatment upon the true endpoint based on the observed effect of treatment on the surrogate.[12] That is, the surrogate needs to be responsive to the treatment and such response needs to translate into an expected response on the clinically relevant endpoint. Therefore, for structural measurements of the optic disc to be considered as reliable surrogate endpoints in glaucoma they need to be

shown to predict future clinically relevant functional outcomes while also being responsive to therapeutic interventions in the disease.

In the current study, we assessed the feasibility of using neuroretinal rim area measurements as surrogate endpoints for development of visual field loss in glaucoma. Using long-term follow-up of a cohort of patients suspected of glaucoma, we evaluated whether progressive changes in neuroretinal rim area were predictive of future development of visual field loss in the disease. In addition, we assessed whether the effect of treatment on the risk of visual field development could be substantially explained by the effect of treatment on neuroretinal rim area losses over time.

METHODS

This was an observational cohort study. Participants from this study were included in two prospective longitudinal studies designed to evaluate optic nerve structure and visual function in glaucoma (the African Descent and Glaucoma Evaluation Study [ADAGES] and the Diagnostic Innovations in Glaucoma Study [DIGS]). The 3-site ADAGES collaboration includes the Hamilton Glaucoma Center and the Department of Ophthalmology, University of California-San Diego (UCSD) (data coordinating center), the New York Eye and Ear Infirmary and the Department of Ophthalmology, University of Alabama at Birmingham, Birmingham (UAB). Although the DIGS includes only patients recruited at UCSD, the protocol of the two studies are identical. The institutional review boards at all 3 sites approved the study methodology, which adhered to the tenets of the Declaration of Helsinki and to the Health Insurance Portability and Accountability Act. All participants gave written informed consent. Methodological details have been described previously.[13]

At each visit during follow-up, subjects underwent a comprehensive ophthalmologic examination including review of medical history, best-corrected visual acuity, slit-lamp biomicroscopy, intraocular pressure (IOP) measurement with Goldman applanation tonometry, gonioscopy, dilated fundoscopic examination, confocal scanning laser ophthalmoscopy and automated perimetry using Swedish Interactive Threshold Algorithm (SITA Standard 24-2). Only subjects with open angles on gonioscopy were included. Subjects were excluded if they presented with a best-corrected visual acuity less than 20/40, spherical refraction outside ± 5.0 diopters and/or cylinder correction outside 3.0 diopters, or any other ocular or systemic disease that could affect the optic nerve or the visual field.

This study included eyes suspected of having glaucoma at the baseline visit. This was based on the presence of suspicious appearance of the optic disc (neuroretinal rim thinning, excavation or suspicious RNFL defects) or elevated IOP (>21mmHg), but normal standard automated perimetry tests at baseline. Normal visual fields were defined based on mean deviation (MD) and pattern standard deviation (PSD) within 95% confidence limits and a Glaucoma Hemifield Test (GHT) within normal limits. All visual fields were evaluated by the UCSD Visual Field Assessment Center (VisFACT).[14] Visual fields with more than 33% fixation losses or false-negative errors, or more than 15% false-positive errors were excluded.

Each patient was required to have a minimum of 5 CSLO examinations per eye during a minimum of two years of follow-up. During follow-up, patients were treated at the discretion of the attending ophthalmologist. Although the treating physician could have done additional testing, the only tests available for analyses were those conducted as part of the study protocol.

Confocal Scanning Laser Ophthalmoscopy

Topographic optic disc measurements were obtained using confocal scanning laser ophthalmoscopy (CSLO) with the HRT II and analyzed using software version 3.0 (Heidelberg Engineering, Dossenheim, Germany). The HRT uses confocal scanning laser principles to obtain a three-dimensional topographic image of the optic nerve. Its principles of operation have been described in detail elsewhere.[15] For each patient, three topographical images were obtained and were combined and automatically aligned to obtain a single mean topography used for analysis. The Imaging Data Evaluation and Analysis (IDEA) Reading Center at the University of California, San Diego conducted all quality assessment and image processing, and certified all operators according to standard protocol.[16] An IDEA Center experienced examiner outlined the optic disc margin on the mean topographic image while viewing stereoscopic photographs of the optic disc. Good quality images required a focused reflectance image with a standard deviation not greater than 50μm. Corneal curvature measurements were used to correct images for magnification error. The parameter "rim area" was used to evaluate optic disc changes over time. This parameter has been described as having the best reproducibility for assessment of longitudinal changes with CSLO.[17]

Visual Field Endpoint

The visual field endpoint was considered as the development of a repeatable abnormal visual field defect on SAP, which required the presence of a sequence of 3 consecutive abnormal tests with PSD with P<5% or Glaucoma Hemifield Test outside normal limits.[18] The date of the first abnormal visual field test was considered as the event date for these eyes. In order to evaluate whether rim area changes were predictive of the development of visual field loss, only CSLO images acquired before the event date were analyzed in the study. Eyes that did not develop the visual field endpoint were considered censored at the last visual field date. All CSLO images up to the last visual field date were analyzed for these eyes.

Joint Longitudinal Survival Model

A joint longitudinal survival model was used to investigate the relationship between longitudinal rim area measurements and risk of visual field development in glaucoma. These models are ideally suited to study the association between changes in a longitudinal marker and the risk for an event, and have been described in detail elsewhere.[19] In brief, they are composed of a longitudinal submodel and a survival submodel which are tied together by sharing random effects. The longitudinal submodel was composed of a linear mixed model with the following formulation:

$$\begin{cases} y_i(t) = m_i(t) + \varepsilon_i(t) \\ m_i(t) = X_i\beta + Z_i b_i \\ b_i - N(0, D), \quad \varepsilon_i(t) \ - N(0, \sigma^2) \end{cases}$$

The model specifically accounts for measurement error of the marker by postulating that the observed level of the outcome $y_i(t)$, corresponding to the CSLO rim area measurements, equals the unobserved true value $m_i(t)$ plus a random error term, $\varepsilon_i(t)$. Covariates can be included in the estimation process of $m_i(t)$ by means of the design matrices X_i and Z_i, for the fixed-effects regression coefficients β, and random-effects regression coefficients b_i, respectively. The random effects were assumed to be normally distributed with mean zero and variance-covariance matrix D. In this particular application, we evaluated the effect of the baseline covariates age, race, corneal thickness and optic disc area, and the follow-up covariates IOP and PSD, on the intercept and slopes of rim area change. The mixed model

assumes random slopes and random intercepts, allowing different rates of change and intercept values for each eye. A hierarchical level can easily be implemented to the model to take into account the dependency between the two eyes of the same subject.

To quantify the strength of the association between the longitudinal marker and the risk for the event (development of visual field loss), a survival submodel was used with the form:

$$h_i(t)=h_0(t)\exp[\gamma_1^{\gamma}w_i+\gamma_2^{\gamma}v_i+\alpha_1 m_i(t_0)+\alpha_2 m_i^{'}(t)],\ \text{where}\ m_i^{'}=\frac{d}{dt}m_i(t),$$

In the survival submodel, $h_i(t)$ determines the hazard function at time *t*, h_0 denotes the baseline hazard function specified by a Weibull distribution, w_i is a vector of baseline covariates with corresponding vector of coefficients (γ_1), v_i is a vector of time-dependent covariates with corresponding vector of coefficients (γ_2). This model was estimated jointly with the longitudinal submodel and allowed an evaluation of the relationship between the true marker values $m_i(t)$ and the risk for the event. We were mainly interested whether the slopes of change in the marker (i.e., rates of change in rim area) were associated with risk of visual field loss. Therefore, $m_i^{'}$ measured the first derivative (slope) of the marker profile and the coefficient $\alpha 2$ measured how strongly associated was the value of the slope of the true longitudinal marker at time t with the risk for an event at the same time point, adjusting for the intercept value and values of other covariates. The interpretation of a is straightforward as in regular survival models, with $\exp(\alpha)$ corresponding to the (HR) ratio for a one unit change in the slope of the marker.

Evaluation of Predictive Ability and Surrogacy

A good surrogate should exhibit a strong correlation with the clinically relevant outcome.[12] Therefore, we evaluate whether changes in rim area over time were strongly predictive of development of visual field loss. To assess the strength of this relationship, we used a survival-adapted adjusted R^2 measure[20] obtained from the model described above. This measure explains the proportion of the outcome variation (i.e, time to development of visual field loss) that can be explained by the predictive factors.[21] A bias-corrected value for R^2 was obtained by 5-fold cross-validation. For cross-validation, the sample was divided into 5 approximately equal and mutually exclusive subsets. Model parameters were obtained from 4 subsets and an R^2 value was obtained by applying the model to the fifth subset. The procedure was repeated 5 times and the final R^2 was the average of the 5 R^2 values obtained from the testing sets.

Another required condition for a useful surrogate is that the effect of the treatment on the surrogate endpoint must reliably predict the effect of treatment on the final clinically relevant endpoint.[12] In the current application, this would mean demonstrating that the effect of treatment on rates of rim area change is a strong predictor of the effect of treatment on the risk of development of visual field loss in glaucoma. This is important because if one wants to use rim area measurements as a surrogate endpoint, it is important to demonstrate that these measurements are responsive to treatment and that this response predicts the response of the final clinically relevant outcome. In order to evaluate this, we calculated the proportion of treatment effect (PTE) explained by the proposed surrogate longitudinal marker (rim area measurements).[22–24] If rim area measurements are a good surrogate for development of visual field loss, they should explain a large proportion of treatment effect. The effect of treatment was established by evaluating the effect of IOP measurements over time by: $PTE=1-\frac{\beta_s}{\beta_0}$, where β_s is the effect of IOP on the risk of development of visual field loss when incorporating the longitudinal marker (rim area) in the same model (i.e., the

survival submodel of the joint model) and β_0 is the treatment effect obtained by simply fitting a survival model evaluating the effect of IOP on the risk of visual field loss excluding the longitudinal marker. Note that if IOP exerts most of its effect on the final outcome (visual field loss) by means of changes in rim area, β_s will tend to be small relative to β_0 and PTE will be close to 1. A high PTE would indicate that the effect of IOP on the risk of visual field loss could be reliably predicted by the effect of IOP on rim area measurements over time. In contrast, if the effect of IOP on the risk of visual field loss cannot be predicted by its effect on rim area measurements over time, β_s will be similar to β_0 and the PTE will be close to zero. A low PTE would indicate that the effect of IOP on the development of field loss would be largely independent of what happens to the rim area measurements. Therefore, the marker would not be a good surrogate to assess the effects of treatment in clinical practice.

Statistical Analyses were performed using STATA v. 12.0 (StataCorp, College Station, Texas). The alpha level (type I error) was set at 0.05.

RESULTS

The study included 492 eyes of 328 patients classified as suspected of having glaucoma at the baseline visit. These eyes had an average of 7.4 ± 2.8 CSLO images during a mean follow-up time of 6.6 ± 1.6 years. Sixty-two (13%) eyes of 44 patients developed visual field loss during follow-up. Table 1 shows baseline clinical and demographic characteristics of eyes that developed versus eyes that did not develop visual field loss during follow-up.

The mean rate of rim area change in eyes that developed visual field loss was $-0.011 mm^2$/year versus $-0.003 mm^2$/year in those that did not (P <0.001). Figure 1 shows the rates of rim area change in the two groups.

Results of the joint longitudinal survival model are shown on Table 2. In the interpretation of the longitudinal submodel (Table 2, top), coefficients for the main effects represent the effect of the covariates on the baseline rim area measurements, whereas terms with interactions with time represent the effect of covariates on changes in rim area over time. In the analysis of factors associated with rates of rim area change during follow-up, each 1mmHg higher IOP during follow-up was associated with a $0.0005 mm^2$/year faster rate of rim area loss over time (P = 0.002). Larger optic disc areas were also associated with faster rates of rim area loss (P = 0.034). Black race was associated with faster rates of rim area loss even after adjustment for the other variables (P = 0.006), while older age was of borderline significance (P = 0.097).

The rate of rim area change was significantly predictive of the risk of development of visual field loss during follow-up. In the univariable model, each $0.01 mm^2$/year faster rate of rim area loss was associated with 2.93 higher risk of development of field loss (HR = 2.93; 95% confidence interval [CI]: 1.67 – 5.16; P<0.001). This corresponded to an R^2 of 62% (95% CI: 53% – 74%). Table 2 (bottom) shows the results of the survival submodel, adjusting for other variables. In the interpretation of the multivariable survival submodel, hazard ratios can be obtained by the exponential of the coefficient value, i.e., exp(coefficient). In the multivariable survival submodel, each $0.01 mm^2$/year faster rate of rim area loss was associated with a 2.94 higher risk of visual field loss (HR = 2.94; 95% CI: 1.38 – 6.23; P = 0.005). The multivariable model had an adjusted R^2 of 81% (95% CI: 71% – 90%) in predicting visual field loss. Figure 2 shows survival probabilities, i.e., probabilities of retaining a normal visual field, for different values of slopes of rim area change over time. It can be seen that eyes with fast slopes of rim area loss had high probabilities of developing field losses during follow-up or, in other words, low probabilities of retaining a normal field (surviving).

From the results of the joint model it was also possible to obtain individual survival probabilities for specific eyes based on results of rim area measurements over time and information on the other risk factors. Figure 3 shows predicted survival probabilities for two eyes, one that showed a relatively fast rate of rim area loss during follow-up (right panel) and another that showed stable measurements over time (left panel). A comparison of the predicted survival probabilities shows that the eye with fast progression had much lower predicted probabilities of survival. This eye in fact showed development of visual field loss during follow-up whereas the eye with stable rim area measurements did not develop any field defect. Figure 4 shows how survival probabilities can be continuously updated during follow-up as more information becomes available. The time course of optic disc changes as seen in stereophotographs and visual field results for the same eye are also shown. The predicted survival probabilities were relatively high when only baseline measurements were considered. As more information became available and a clear trend of rim area loss was observed, the model estimated much lower probabilities of survival. The results of rim area assessment were in agreement with changes observed on optic disc photographs and the eye later developed a visual field defect.

When the effect of rim area changes was considered in the joint model (Table 2), IOP was not significantly associated with risk of development of visual field loss (HR = 1.03 per 1mmHg higher; 95% CI: 0.94 – 1.13; P = 0.548). This was in contrast to the survival model that excluded longitudinal rim area measurements (Table 3). In the absence of the effect of rim area changes over time in the model, IOP was significantly associated with risk of development of visual field loss with HR of 1.09 per 1mmHg higher (95% CI: 1.01 – 1.17; P=0.029). From the coefficients representing the effect of IOP on risk of visual field loss in the joint model ($\beta_s = 0.0287$) (Table 2) and the model without longitudinal rim area changes ($\beta_0 = 0.0816$), the PTE was calculated as $\beta_0/\beta s = 1 - 0.0287/0.0816 = 0.65$ or 65%.

DISCUSSION

In the current study, we demonstrated that progressive neuroretinal rim loss was predictive of development of visual field loss in a cohort of glaucoma suspects followed over time. The use of a joint longitudinal survival model allowed us to quantify the ability of rates of rim area change in predicting the risk of functional loss adjusting for the effect of confounding variables. In addition, we evaluated whether the effect of treatment as measured by IOP on rates of rim area change was predictive of the effect of treatment on the risk of development of visual field loss in glaucoma, a necessary condition for a suitable surrogate endpoint. To our knowledge, this is the first study to evaluate the potential of a biomarker as a surrogate endpoint in glaucoma. Our findings may have significant implications for the use of CSLO rim area measurements in clinical practice and also in clinical trials evaluating glaucoma progression with this technology.

A validated surrogate endpoint is an endpoint that allows prediction of a clinically relevant endpoint but in itself does not measure a clinical benefit.[2] Therefore, although structural measurements of the optic nerve do not directly measure a clinical benefit for the patient, if they can be shown to predict a clinically relevant outcome, they could be considered as potentially useful surrogates. In a landmark study, Prentice[12] formulated a set of operational criteria for validating a surrogate endpoint: (I) The surrogate endpoint has a significant impact on the true endpoint; (II) treatment has a significant impact on the surrogate endpoint; (III) treatment has a significant impact on the true endpoint and (IV) the full effect of treatment upon the true endpoint can be captured by the surrogate. Therefore, a first step in validating a proposed surrogate is to demonstrate that it can reliably predict the clinically relevant outcome (Criterion I). In our study, we showed that progressive rim loss was predictive of development of visual field loss. In the multivariable model, each $0.01mm^2$/

year faster rate of rim area loss was associated with a 2.94 higher risk of visual field loss (HR = 2.94; 95% CI: 1.38 – 6.23; P = 0.005). More importantly, when the predictive ability was quantified, the model including only slope and intercept of rim area measurements was able to explain 62% of the variation in the clinically relevant outcome, i.e., time to development of visual field loss. When other covariates were included, the adjusted R^2 increased to 81%. These results agree with the study by Medeiros et al[10] showing that progressive optic disc changes seen on optic disc stereophotographs are highly predictive of functional losses in glaucoma. A previous study by Chauhan et al[25] also evaluated whether rim area changes measured by the topographic change analysis (TCA) algorithm were predictive of visual field progression in glaucoma. The authors found that the TCA was able to predict future progression of visual field loss. However, their analysis did not take into account the effect of potentially confounding variables, such as intraocular pressure, corneal thickness or visual field measurements. In addition, they did not use statistical methods to deal with the censored aspect of this type of data, which may result in biased estimates of significance of predictive factors and predictive ability.[26] In our study, we used a joint longitudinal survival model that allows us to evaluate the predictive ability of a proposed biomarker, taking into account measurement error and also adjusting for the confounding effect of other relevant variables. In our analysis, rim area changes were predictive of future development of visual field loss even when other variables previously shown to be associated with risk of progression were included in the model. Importantly, the model adjusted for PSD measurements acquired over time. This shows that rim area changes provided additional information besides what could be obtained from visual field analysis in predicting the future functional outcome.

Although rim area changes were highly predictive of risk of development of visual field loss, such finding is not sufficient for validating these measurements as surrogate endpoints. As Fleming and DeMets[27] pointed out “a correlate does not make a surrogate”. In fact, Prentice criteria also require demonstrating an effect of treatment on the surrogate (criterion II) and true endpoint (criterion III), as well as demonstrating that the effect of the treatment on the surrogate can reliably predict the effect on the true endpoint (criterion IV). In our analysis, we used IOP measurements over time as indicators of the effect of treatment. We showed that each 1mmHg higher IOP was associated with a 0.0005mm^2/year faster rate of rim area loss over time (P = 0.002). The significant effect of treatment on the proposed surrogate seems to satisfy Prentice’s criterion II. We also showed that IOP had a significant effect on the development of visual field loss in the model not including rim area measurements, as seen on Table 3, which satisfies Prentice’s criterion III. In order to demonstrate Prentice criterion IV, that the full effect of treatment upon the true endpoint can be captured by the surrogate, one requires a statistical test for the treatment effect on the true endpoint to be nonsignificant after adjustment for the surrogate.[12] The results of the joint longitudinal survival model (Table 2) show that the effect of IOP in predicting the risk of development of visual field loss was nonsignificant after adjustment for longitudinal rim area measurements (P = 0.548). Several authors, however, have criticized this approach as the lack of significance of the treatment effect in the adjusted model could be just due to a lack of power from inadequate sample size.[22] Therefore, it has been proposed that the PTE explained by the surrogate would be a better method of quantification and validation of the proposed surrogate.[13, 22] A high PTE would indicate that the effect of treatment on the true outcome would be reliably predicted by the effect of treatment on the surrogate. This is an essential condition as reliance on surrogate outcomes is justifiable only if treatment comparisons that are based on the surrogate are a faithful reflection of comparisons that are based on the true clinically relevant endpoint. We found a PTE of 65% for longitudinal rim area measurements. That is, rim area measurements were able to explain 65% of the effect of treatment on the risk of development of visual field loss. Although this effect can be considered only moderate, it should be noted that a PTE of 100% is rarely, if ever, seem in

clinical practice.[24, 28] Further, the use of rim area measurements as endpoints in clinical trials does not necessarily preclude the concomitant use of visual fields for detecting clinically relevant endpoints. In fact, a composite endpoint combining structure and function has been shown to detect larger number of cases of disease progression than the isolated use of these methods.[29–32] Therefore, the inclusion of visual field assessment along with structural testing would allow a more complete evaluation of treatment effect, while still retaining the ability of earlier identification of relevant endpoints.

Variations of IOP have been shown to influence CSLO topographic optic disc measurements in a way that does not reflect the effect of IOP in causing progressive neural tissue loss in glaucoma. For example, improvements in rim area measurements are seen from IOP reduction after glaucoma surgery or ocular hypotensive treatment but such improvements do not translate into improvement in visual function.[33] This can result in a decrease in the ability of topographic optic disc measurements in capturing the effect of IOP on the clinically relevant functional outcome, reducing the PTE and the utility of these measurements as surrogate endpoints. Therefore, if this effect were to be removed, the PTE for rim area could potentially be even higher. This confounding effect seems to be less pronounced on structural measurements obtained by other technologies, such as retinal nerve fiber layer assessment with optical coherence tomography. Therefore, it is possible that retinal nerve fiber layer measurements may be able to capture a larger proportion of treatment effect on relevant clinical outcomes compared to topographic optic disc measurements. Further studies are necessary to evaluate this hypothesis.

The joint longitudinal survival model presented in our study also allowed estimation of individual survival probabilities over time. Using this model, the risk of development of visual field loss can be updated as information on predictive factors is made continuously available over time, as shown on Figure 4. Such approach offers significant advantages over currently available predictive models or risk calculators designed to estimate risk of glaucoma development.[34, 35] These models use only baseline information on predictive factors, which results in limited value in predicting outcomes. In fact, analysis of the predictive model containing only baseline information in our study revealed an R^2 of only 33% (95% CI: 18% – 54%), compared to 81% for the model including longitudinal information. This is in line with the results of the CSLO ancillary study to the Ocular Hypertension Treatment Study. In that study, baseline CSLO measurements were found to be predictive of future glaucoma conversion in ocular hypertension eyes.[36] However, the predictive ability of baseline CSLO measurements was actually relatively poor. In our study, we showed a much better predictive ability when longitudinal measurements were taken into account rather than just baseline information. As Figure 4 illustrates, the probabilities of progression can be continuously updated as more information becomes available, resulting in more effective use of clinical information. Similarly, the two eyes shown on Figure 3 had similar baseline CSLO measurements, but their risks of progression were very different when longitudinal information was incorporated into the model. Using the model presented in our study, a glaucoma suspect that is longitudinally followed over time can have its predicted probability of developing functional loss readjusted over time, as more imaging measurements and information on risk factors become available. Survival curves with 95% confidence bands like those of Figure 4 could be provided to clinicians and decisions about treatment could be based on the estimated probabilities of functional loss, taking also into account considerations about life expectancy and risks of treatment. Future studies should evaluate how to best integrate all this information into clinical decision-making about treatment of glaucoma suspects.

Our study has several limitations. We used development of visual field loss as indicative of the true clinically relevant endpoint. It might be arguable that a true endpoint would

represent measurements more directly related to quality of life, such as patient-reported outcomes or results of performance-based tests. However, recent analysis of population-based data has suggested that even mild visual field loss in glaucoma patients already carries a significant negative impact in vision-related quality of life measures.[37] Other studies have shown that because of the nonlinear relationship between retinal ganglion cell loss and perimetric measurements, early visual field loss may already be associated with substantial loss of retinal ganglion cells and that even relatively smaller additional losses could potentially lead to severe field defects.[38, 39] Despite these observations, it will be important to conduct further longitudinal studies validating structural measurements as surrogate endpoints for quality of life-based outcomes in glaucoma. Another limitation of our study is that it did not randomly assign subjects to treatment versus no treatment. Although we did not employ a randomization protocol, we used longitudinal IOP measurements as estimates of treatment effect, which is a reasonable assumption for the relevant treatments currently available for glaucoma. However, it is important to exercise caution when extrapolating the results of the present investigation to the context of a clinical trial validating non-IOP-related therapies, such as potential neuroprotective agents.

In conclusion, progressive neuroretinal rim area changes were highly predictive of future development of visual field losses in glaucoma and explained a significant proportion of treatment effect on the clinically relevant outcome. These findings suggest that rim area measurements may be suitable surrogate endpoints in glaucoma clinical trials. The use of joint longitudinal survival models can also provide individualized risk predictions that may be superior to currently existing approaches for estimation of risk of progressive damage in the disease.

Acknowledgments

Supported in part by NIH/NEI grants EY021818 (FAM), EY11008 (LMZ), EY14267 (LMZ), CAPES grant BEX 1066/11-0 (RL), an unrestricted grant from Research to Prevent Blindness (New York, NY), grant for participants' glaucoma medications from Alcon, Allergan, Pfizer, Merck and Santen.

REFERENCES

1. Medeiros, FA.; Susanna, R., Jr; Singh, K. Who should be treated?. In: Weinreb, RN.; Liebmann, J., editors. Medical Treatment of Glaucoma. Amsterdam, The Netherlands: Kugler; 2010. p. 1-19.
2. Biomarkers Definitions Working Group. Biomarkers and surrogate endpoints: preferred definitions and conceptual framework. Clin Pharmacol Ther. 2001; 69:89–95. [PubMed: 11240971]
3. Boissel JP, Collet JP, Moleur P, Haugh M. Surrogate endpoints: a basis for a rational approach. Eur J Clin Pharmacol. 1992; 43:235–244. [PubMed: 1425885]
4. Kass MA, Heuer DK, Higginbotham EJ, et al. Ocular Hypertension Treatment Study Group The Ocular Hypertension Treatment Study: a randomized trial determines that topical ocular hypotensive medication delays or prevents the onset of primary open-angle glaucoma. Arch Ophthalmol. 2002; 120:701–713. discussion 829–30. [PubMed: 12049574]
5. Drance SM. The Collaborative Normal-Tension Glaucoma Study and some of its lessons [in English, French]. Can J Ophthalmol. 1999; 34:1–6. [PubMed: 10088056]
6. Katz J, Congdon N, Friedman DS. Methodological variations in estimating apparent progressive visual field loss in clinical trials of glaucoma treatment. Arch Ophthalmol. 1999; 117:1137–1142. [PubMed: 10496384]
7. Kamal DS, Garway-Heath DF, Hitchings RA, Fitzke FW. Use of sequential Heidelberg retina tomograph images to identify changes at the optic disc in ocular hypertensive patients at risk of developing glaucoma. Br J Ophthalmol. 2000; 84:993–998. [PubMed: 10966952]
8. Strouthidis NG, Scott A, Peter NM, Garway-Heath DF. Optic disc and visual field progression in ocular hypertensive subjects: detection rates, specificity, and agreement. Invest Ophthalmol Vis Sci. 2006; 47:2904–2910. [PubMed: 16799032]

9. Medeiros FA, Zangwill LM, Bowd C, et al. Use of progressive glaucomatous optic disk change as the reference standard for evaluation of diagnostic tests in glaucoma. Am J Ophthalmol. 2005; 139:1010–1018. [PubMed: 15953430]
10. Medeiros FA, Alencar LM, Zangwill LM, et al. Prediction of functional loss in glaucoma from progressive optic disc damage. Arch Ophthalmol. 2009; 127:1250–1256. [PubMed: 19822839]
11. Fleming TR, DeMets DL. Surrogate end points in clinical trials: are we being misled? Ann Intern Med. 1996; 125:605–613. [PubMed: 8815760]
12. Prentice RL. Surrogate endpoints in clinical trials: definition and operational criteria. Stat Med. 1989; 8:431–440. [PubMed: 2727467]
13. Sample PA, Girkin CA, Zangwill LM, ADAGES Study Group, et al. The African Descent and Glaucoma Evaluation Study (ADAGES): design and baseline data. Arch Ophthalmol. 2009; 127:1136–1145. [PubMed: 19752422]
14. Racette L, Liebmann JM, Girkin CA, et al. ADAGES Group. African Descent Glaucoma Evaluation Study (ADAGES): III Ancestry differences in visual function in healthy eyes. Arch Ophthalmol. 2010; 128:551–559. [PubMed: 20457975]
15. Weinreb RN. Laser scanning tomography to diagnose and monitor glaucoma. Curr Opin Ophthalmol. 1993; 4:3–6. [PubMed: 10148455]
16. Girkin CA, Sample PA, Liebmann JM, et al. ADAGES Group. African Descent Glaucoma Evaluation Study (ADAGES): II Ancestry differences in optic disc, retinal nerve fiber layer, and macular structure in healthy subjects. Arch Ophthalmol. 2010; 128:541–550. [PubMed: 20457974]
17. Asaoka R, Strouthidis NG, Kappou V, et al. HRT-3 Moorfields reference plane: effect on rim area repeatability and identification of progression. Br J Ophthalmol. 2009; 93:1510–1513. [PubMed: 19535359]
18. Keltner JL, Johnson CA, Levine RA, et al. Ocular Hypertension Treatment Study Group Normal visual field test results following glaucomatous visual field end points in the Ocular Hypertension Treatment Study. Arch Ophthalmol. 2005; 123:1201–1206. [PubMed: 16157799]
19. Wulfsohn MS, Tsiatis AA. A joint model for survival and longitudinal data measured with error. Biometrics. 1997; 53:330–339. [PubMed: 9147598]
20. Royston P, Sauerbrei W. A new measure of prognostic separation in survival data. Stat Med. 2004; 23:723–748. [PubMed: 14981672]
21. Choodari-Oskooei B, Royston P, Parmar MK. A simulation study of predictive ability measures in a survival model I: explained variation measures. Stat Med. 2012; 31:2627–2643. [PubMed: 21520455]
22. Lin DY, Fleming TR, De Gruttola V. Estimating the proportion of treatment effect explained by a surrogate marker. Stat Med. 1997; 16:1515–1527. [PubMed: 9249922]
23. Flandre P, Saidi Y. Estimating the proportion of treatment effect explained by a surrogate marker [letter]. Stat Med. 1999; 18:107–109. [PubMed: 9990696]
24. Freedman LS, Graubard BI, Schatzkin A. Statistical validation of intermediate endpoints for chronic diseases. Stat Med. 1992; 11:167–178. [PubMed: 1579756]
25. Chauhan BC, Nicolela MT, Artes PH. Incidence and rates of visual field progression after longitudinally measured optic disc change in glaucoma. Ophthalmology. 2009; 116:2110–2118. [PubMed: 19500850]
26. Medeiros FA, Weinreb RN. Visual field progression [letter]. Ophthalmology. 2010; 117:851–852. author reply 852. [PubMed: 20346830]
27. Fleming TR, DeMets DL. Surrogate end points in clinical trials: are we being misled? Ann Intern Med. 1996; 125:605–613. [PubMed: 8815760]
28. O'Brien WA, Hartigan PM, Daar ES, et al. VA Cooperative Study Group on AIDS Changes in plasma HIV RNA levels and CD4+ lymphocyte counts predict both response to antiretroviral therapy and therapeutic failure. Ann Intern Med. 1997; 126:939–945. [PubMed: 9182470]
29. 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. [PubMed: 22840484]
30. Medeiros FA, Zangwill LM, Girkin CA, et al. Combining structural and functional measurements to improve estimates of rates of glaucomatous progression. Am J Ophthalmol. 2012; 153:1197–1205. [PubMed: 22317914]

31. Medeiros FA, Leite MT, Zangwill LM, Weinreb RN. Combining structural and functional measurements to improve detection of glaucoma progression using Bayesian hierarchical models. Invest Ophthalmol Vis Sci. 2011; 52:5794–5803. [PubMed: 21693614]
32. Russell RA, Malik R, Chauhan BC, et al. Improved estimates of visual field progression using Bayesian linear regression to integrate structural information in patients with ocular hypertension. Invest Ophthalmol Vis Sci. 2012; 53:2760–2769. [PubMed: 22467579]
33. Irak I, Zangwill L, Garden V, et al. Change in optic disk topography after trabeculectomy. Am J Ophthalmol. 1996; 122:690–695. [PubMed: 8909209]
34. Medeiros FA, Weinreb RN, Sample PA, et al. Validation of a predictive model to estimate the risk of conversion from ocular hypertension to glaucoma. Arch Ophthalmol. 2005; 123:1351–1360. [PubMed: 16219726]
35. Ocular Hypertension Treatment Study Group, European Glaucoma Prevention Study Group. Validated prediction model for the development of primary open-angle glaucoma in individuals with ocular hypertension. Ophthalmology. 2007; 114:10–19. [PubMed: 17095090]
36. Zangwill LM, Weinreb RN, Beiser JA, et al. Baseline topographic optic disc measurements are associated with the development of primary open-angle glaucoma: the Confocal Scanning Laser Ophthalmoscopy Ancillary Study to the Ocular Hypertension Treatment Study. Arch Ophthalmol. 2005; 123:1188–1197. [PubMed: 16157798]
37. McKean-Cowdin R, Varma R, Wu J, et al. Los Angeles Latino Eye Study Group Severity of visual field loss and health-related quality of life. Am J Ophthalmol. 2007; 143:1013–1023. [PubMed: 17399676]
38. Medeiros FA, Lisboa R, Weinreb RN, et al. Retinal ganglion cell count estimates associated with early development of visual field defects in glaucoma. Ophthalmology. 2013; 120:736–744. [PubMed: 23246120]
39. Medeiros FA, Zangwill LM, Bowd C, et al. The structure and function relationship in glaucoma: implications for detection of progression and measurement of rates of change. Invest Ophthalmol Vis Sci. 2012; 53:6939–6946. [PubMed: 22893677]

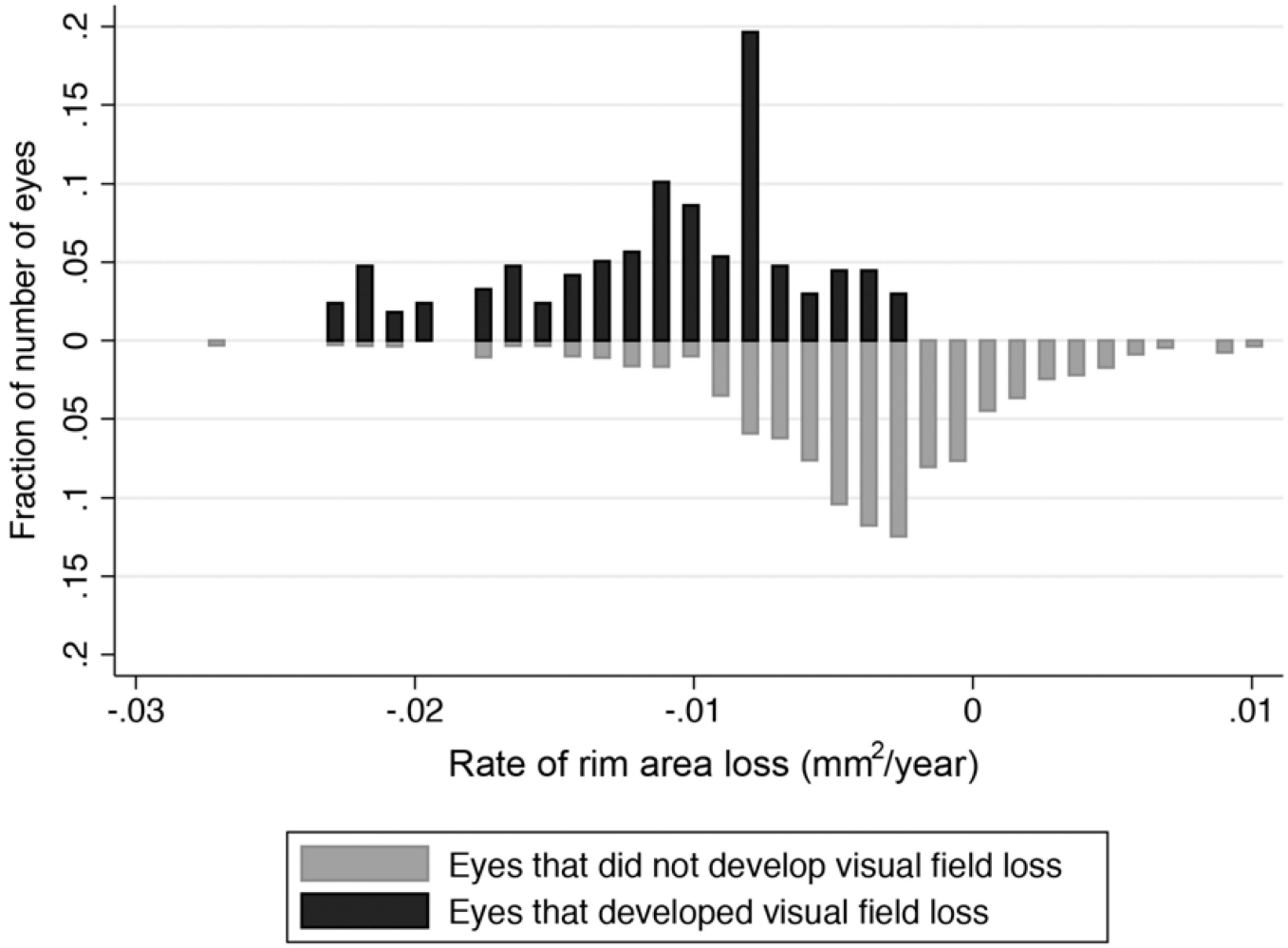

Figure 1.
Distribution of rates of neuroretinal rim area loss in eyes that developed visual field loss versus eyes that did not develop visual field loss.

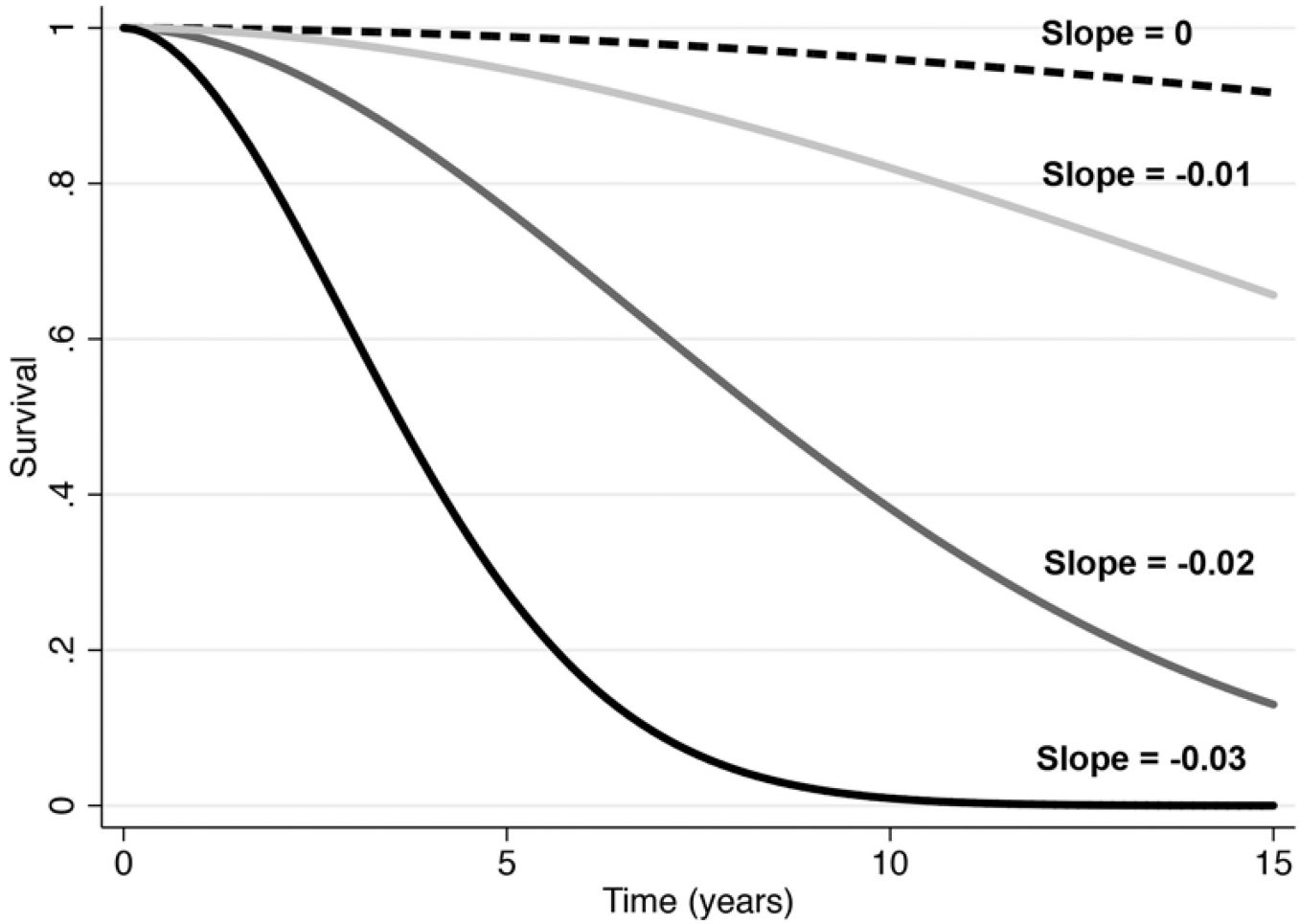

Figure 2.
Survival, i.e., probabilities of retaining a normal visual field, during follow-up for different values of slopes (rates) of rim area loss over time. Slopes are given as mm^2/year. Eyes with faster slopes of rim area loss (more negative values) had lower probabilities of survival over time.

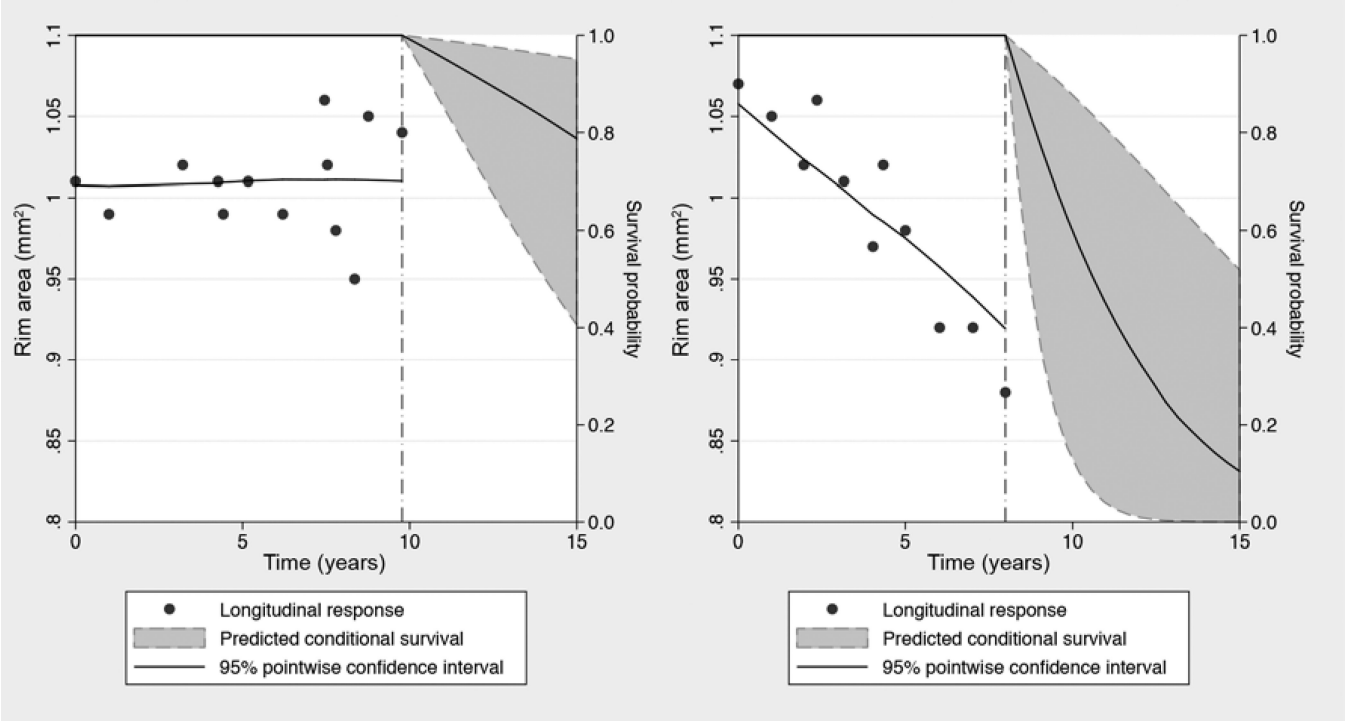

Figure 3.

Predicted survival probabilities for two eyes, one that showed a relatively fast rate of rim area loss during follow-up (right panel) and another that showed stable measurements over time (left panel). A comparison of the predicted survival probabilities shows that the eye with fast progression had much lower predicted probabilities of survival, i.e., retaining a normal visual field. This eye in fact showed development of visual field loss during follow-up whereas the eye with stable rim area measurements did not develop any field defect.

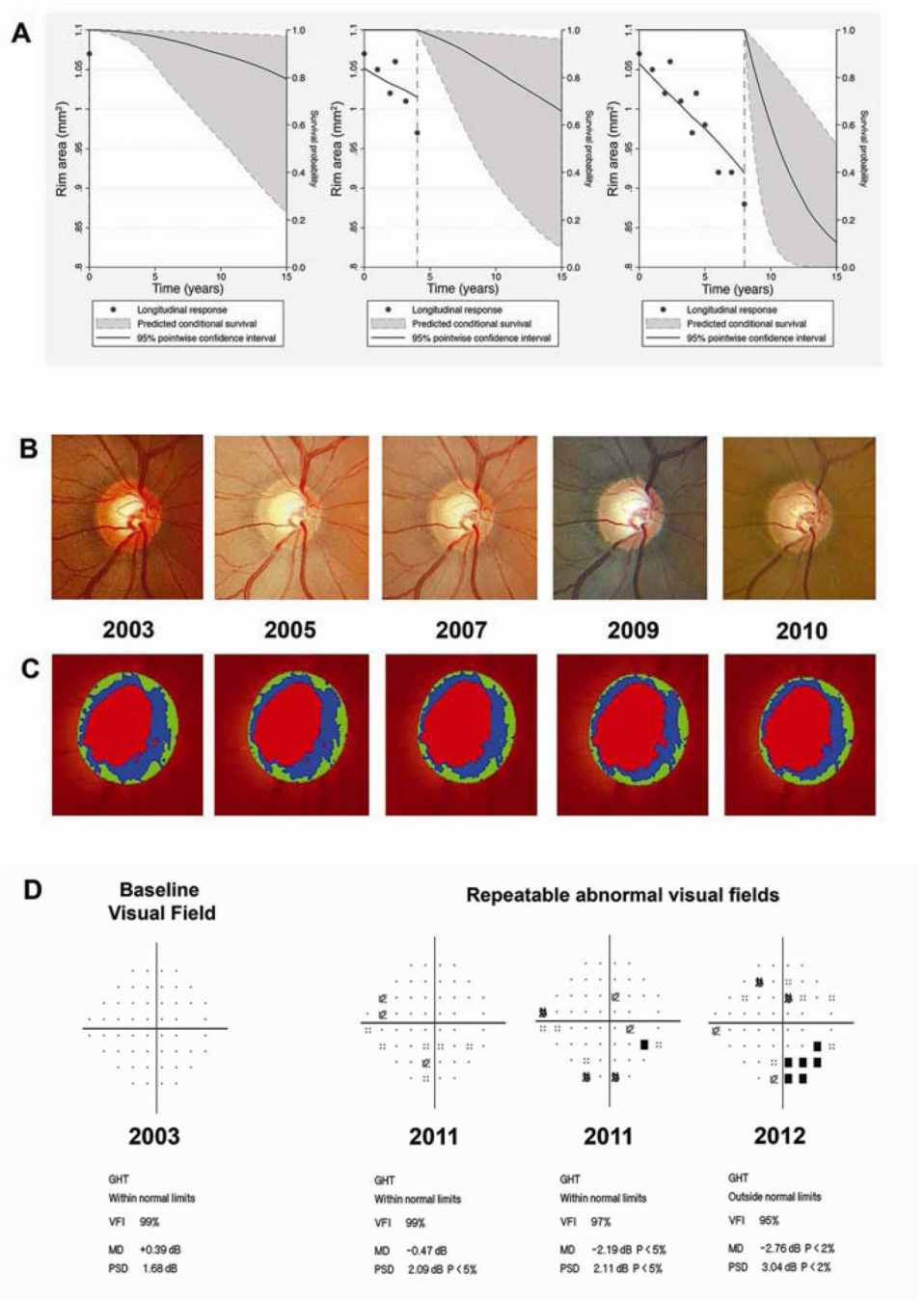

Figure 4.

Example of how survival probabilities can be updated as more information on predictive factors becomes available during follow-up. **A**. Left panel shows survival probabilities after considering only the baseline data. The model estimated that the probability of retaining a normal visual field over time was relatively high. As more information became available (middle and right panels), the survival probabilities were updated. The estimated survival probabilities became much lower as the result of progressive rim area losses over time. **B**. Corresponding optic disc photographs showing progressive loss of neuroretinal rim tissue. **C**. Corresponding confocal scanning laser ophthalmoscopy topographic images showing

progressive loss of rim area over time. **D**. This eye developed a repeatable visual field defect at the end of follow-up.
GHT = Glaucoma Hemifield Test; VFI = visual field index; MD= mean deviation; PSD = pattern standard deviation

Table 1

Demographic and clinical characteristics of eyes that developed and eyes that did not develop visual field loss during follow-up.

Parameter	Developed visual field loss (n = 62 eyes, 44 patients)	Did not develop visual field loss (n = 430 eyes, 284 patients)	P
Age at baseline years	58.3 ± 12.9	56.6 ± 11.6	0.265
Gender, % female	59%	62%	0.744
Race			
White, n (%)	20 (41)	195 (70)	
Black, n (%)	29 (59)	84 (30)	<0.001
Mean intraocular pressure, mmHg	19.8 ± 5.6	17.8 ± 3.7	<0.001
Central corneal thickness, μm	545.2 ± 45.9	555.4 ± 38.7	0.059
Baseline mean deviation, dB	−1.12 ± 1.16	−0.18 ± 1.29	<0.001
Baseline pattern standard deviation dB	1.66 ± 0.22	1.46 ± 0.24	<0.001
Rim area at baseline mm^2	1.32 ± 0.37	1.35 ± 0.27	0.532
Optic disc area, mm^2	2.13 ± 0.43	2.10 ± 0.50	0.610
Follow-up time years	4.2 ± 2.2	6.9 ± 1.2	<0.001

* Values are given as mean ± standard deviation, unless otherwise noted.

Table 2

Results of the joint longitudinal survival model investigating the effect of longitudinal changes in rim area in predicting the risk of development of visual field loss, while adjusting for confounding factors.

Longitudinal submodel				
Parameter	**Coefficient**	**95% confidence interval**	**P**	
Constant	1.194	1.08 – 1.31	<0.001	
Age, per decade older	0.022	0.004 – 0.041	0.020	
Race, Black	0.042	−0.008 – 0.092	0.097	
IOP, per 1mmHg higher	0.004	0.001 – 0.008	0.017	
CCT, per 100μm thinner	−0.106	−0.165 – −0.048	<0.001	
Disc Area, per 0.1mm^2 larger	0.026	0.022 – 0.031	<0.001	
Time	0.001	−0.004 – 0.006	0.714	
Time x Age	−0.0007	−0.0017 – 0.0001	0.097	
Time x Race	−0.0034	−0.0058 - −0.001	0.006	
Time x IOP	−0.0005	−0.0008 - −0.0002	0.002	
Time x CCT	−0.002	−0.005 – 0.0006	0.178	
Time x Disc Area	−0.00024	−0.0005 - −0.00002	0.034	
Survival submodel				
Parameter	**Coefficient**	**95% confidence interval**	**P**	**Hazard Ratio**
Slope, per −0.01mm^2/year lower	1.08	0.32 – 1.83	0.005	2.94
Intercept, per 0.1mm^2 higher	0.066	−0.074 – 0.208	0.357	1.07
Age, per decade older	0.06	−0.20 – 0.32	0.644	1.06
Race, Black	0.42	−0.22 – 1.06	0.197	1.52
IOP, per 1mmHg higher	0.0287	−0.0650 – 0.1224	0.548	1.03
CCT, per 100μm thinner	0.85	0.04 – 1.66	0.039	2.34
PSD, per 0.1dB higher	0.13	0.10 – 0.16	<0.001	1.14
Disc Area, per 0.1mm^2 larger	−0.036	−0.106 – 0.034	0.316	0.96

IOP – intraocular pressure; CCT – central corneal thickness; PSD – pattern standard deviation

Table 3

Results of the survival model predicting risk of development of visual field loss excluding longitudinal rim area measurements

Parameter	Coefficient	95% CI	P	Hazard ratio
Age, per decade older	0.008	−0.226 – 0.243	0.944	1.01
Race, Black	0.75	0.19 – 1.31	0.009	2.12
IOP, per 1mmHg higher	0.0816	0.0082 – 0.1550	0.029	1.09
CCT, per 100μm thinner	0.98	0.24 – 1.73	0.009	2.66
PSD, per 0.1dB higher	0.12	0.10 – 0.14	<0.001	1.13
Disc Area, per 0.1mm^2 larger	0.005	−0.046 – 0.056	0.849	1.01

IOP – intraocular pressure; CCT – central corneal thickness; PSD – pattern standard deviation; CI – confidence interval