Visual Field Mean Deviation Slopes: A Potential Primary Endpoint for Clinical Trials in Glaucoma Neuroprotection

Carlos Gustavo De Moraes (cvd2109@columbia.edu)
Columbia University Irving Medical Center

Keith J. Lane
Ora Clinical, Inc

Xiao Wang
Statistics and Data Corporation, Inc

Jeffrey M. Liebmann
Columbia University Irving Medical Center

Article

Keywords: visual field, glaucoma, progression, clinical trials

Posted Date: November 11th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-2195764/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License
Abstract

The purpose of this retrospective, longitudinal study is to evaluate the relationship between MD slope from visual field tests collected over a short period of time (2 years) and the current United States' Food and Drug Administration (FDA) recommended endpoints for visual field outcomes. If this correlation is strong and highly predictive, clinical trials employing MD slopes as primary endpoints could be employed in neuroprotection clinical trials with shorter duration and help expedite the development of novel IOP-independent therapies. Visual field tests of patients with or suspected glaucoma were selected from an academic institution and evaluated based on two functional progression endpoints: 1) five or more locations worsening by at least 7 dB, and 2) at least 5 test locations based upon the GCP algorithm. A total of 467 (76.6%) and 472 (77.4%) eyes reached Endpoints A and B, respectively, at any point in time. 359 (58.8%) eyes progressed faster than 0 dB/year, 278 (45.5%) faster than -0.5 dB/year, and 200 (32.7%) faster than -1.0 dB/year. It was found that eyes experiencing rapid 24-2 visual field MD slopes over a 2-year period were more likely to reach one of the FDA accepted endpoints during or soon after that period.

Introduction

Intraocular pressure (IOP) reduction significantly slows the rate of visual field progression for all glaucoma subtypes and across the disease severity spectrum. However, many patients continue to progress despite significant IOP reduction, and this is often attributed to either pressure-independent mechanisms of retinal ganglion cell (RGC) death, insufficient IOP reduction, or a combination of both. Nonetheless, only IOP lowering has been proven to be a modifiable risk factor for glaucoma onset and progression.

Given recent advances in our understanding of the mechanisms of RGC death, a variety of novel approaches for glaucoma neuroprotection are being investigated. One main challenge for the development of these IOP-independent neuroprotective therapies is the lack of glaucoma progression endpoints that could be tested in clinical trials of relatively short duration in a glaucoma population treated with concomitant IOP-lowering therapies. Most clinical studies to date assessing glaucoma progression endpoints have included large populations followed for 4 or more years; the cost of similar trials would be prohibitive for both neuroprotection proof-of-concept studies and for longer and larger clinical trials aiming for regulatory approval. Moreover, given the high efficacy of currently available glaucoma medications and surgical interventions (both laser and incisional) in slowing the rate of visual field progression, it would not be possible to run such studies on a placebo group of untreated patients. Therefore, the likelihood of finding statistically and clinically meaningful differences between study arms in neuroprotection glaucoma trials would likely be small and would come with substantial cost.

To address these challenges, there have been two joint discussions between the National Eye Institute (NEI) of the National Institutes of Health (NIH) and the Food and Drug Administration Center for Drug Development and Research (FDA CDER) promoted by the Association for Research in Vision and Ophthalmology (ARVO) to discuss glaucoma progression endpoints. The two functional endpoints
Currently accepted by the FDA and discussed during the NEI/FDA Glaucoma Clinical Trial Design and Endpoints Symposium are: (A) visual field progression defined as the between group mean difference in threshold for 5 or more pre-selected visual field locations that are statistically significant and have a difference that is at least 7 dB on more than one examination, and (B) visual field progression defined as five or more reproducible points of visual field locations with significant changes (p < 0.05) from baseline beyond the 5% probability levels for the Glaucoma-Change-Probability (GCP) analysis.\textsuperscript{12,14} Given the high specificity of these criteria, identifying a population of glaucoma patients with a high proportion of subjects reaching such endpoints over a short study period (e.g. 2 years or less) is challenging and would require hundreds or thousands of patients per study arm. Per the committee's discussion, an alternative glaucoma endpoint for pivotal trials should have a strong correlation and predictability of either current or future visual function, such as that measured with the above cited acceptable endpoints.\textsuperscript{12}

Investigators have reported the advantages of using visual field mean deviation (MD) slopes as an endpoint in glaucoma clinical trials because studies employing trend analysis of visual field global indices would likely require a smaller sample size than event-based endpoints,\textsuperscript{15–19} such as the ones currently acceptable by the FDA. The purpose of the present study is to evaluate the relationship between MD slope from visual field tests collected over a short period of time (2 years) and the current FDA recommended endpoints. If this correlation is strong and highly predictive, as the agency recommends, clinical trials employing MD slopes as primary endpoints could be employed in neuroprotection clinical trials with shorter duration and help expedite the development of novel IOP-independent therapies.

### Results

#### Database Characteristics

A total of 9,595 reliable visual field tests from 678 eyes from 558 patients met the main inclusion and exclusion criteria for analyses. Of those, 68 eyes were excluded due to any suspicion of non-glaucomatous causes that may have affected changes in visual function, leaving a final database of 610 eyes of 511 patients. The mean (IQR) number of visual field tests per eye was 14.15 (7 to 36) spanning 6.74 (0.89 to 15.56) years. A total of 467 (76.6%) and 472 (77.4%) eyes reached Endpoints A and B, respectively, at any point in time. 359 (58.8%) eyes progressed faster than 0 dB/year, 278 (45.5%), faster than − 0.5 dB/year, and 200 (32.7%) faster − 1.0 dB/year. Clinical characteristics of the final sample are shown in Table 1. Note that the sample includes a sample of patients who meet the known age demographics for established glaucoma with moderate functional damage and who performed reliable visual field tests during the study.
Table 1
Patient demographics for those individuals meeting the study entry criteria.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Mean</th>
<th>Standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>64.8</td>
<td>14.0</td>
</tr>
<tr>
<td>Baseline MD (dB)</td>
<td>-7.5</td>
<td>3.1</td>
</tr>
<tr>
<td>MD slope (dB/year)</td>
<td>-0.37</td>
<td>1.6</td>
</tr>
<tr>
<td>False-negative responses (%)</td>
<td>5.5</td>
<td>5.5</td>
</tr>
<tr>
<td>False-positive responses (%)</td>
<td>2.6</td>
<td>2.9</td>
</tr>
<tr>
<td>Fixation losses (%)</td>
<td>7.6</td>
<td>8.3</td>
</tr>
</tbody>
</table>

Figure 1 shows the Kaplan-Meier curves for Endpoint A. Each line corresponds to one of the three MD slope cut-off values investigated in addition to the entire final sample. The time range between 0 and 2 years corresponds to Sequence 1. For MD slopes faster than 0 dB/year, the time required to reach an endpoint in 25, 50, and 75% of eyes was 0.9, 1.7, and 3.3 years, respectively (hazard ratio, HR = 2.11, 95% CI = 1.74 to 2.56, P < 0.001). For MD slopes faster than −0.5 dB/year, the time taken was 0.8, 1.4, and 2.4 years, respectively (HR = 2.59, 95% CI = 2.14 to 3.12, P < 0.001). For MD slopes faster than −1.0 dB/year, the time taken was 0.7, 1.2, and 2.1 years, respectively (HR = 2.58, 95% CI = 2.13 to 3.12, P < 0.001).

A similar depiction is shown in Fig. 2 for Endpoint B. For MD slopes faster than 0 dB/year, the time required to reach an endpoint in 25, 50, and 75% of eyes to reach the endpoint was 0.89, 1.65, and 3.29 years, respectively (HR = 2.36, 95% CI = 1.94 to 2.88, P < 0.001). For MD slopes faster than −0.5 dB/year, the time required to reach an endpoint was 0.83, 1.39, and 2.12 years, respectively (HR = 2.76, 95% CI = 2.29 to 3.34, P < 0.001). For MD slopes faster than −1.0 dB/year, the time taken was 0.76, 1.28, and 2.01 years, respectively (HR = 2.74, 95% CI = 2.26 to 3.32, P < 0.001).

To better understand the characteristics of test locations more likely to reach an endpoint during follow up, we investigated their average distribution and baseline severity. Their median (IQR) baseline values were −6.5 dB (-4.0 to -11.0) and mean (SD) were −7.8 dB (5.9). We also looked at each of the 610 eyes to assess a pattern that could describe the relationship between baseline pointwise severity and losing sensitivity according to Endpoints A and B. We found that in most cases the progressing locations (according to these endpoints) were within or adjacent to test locations with significant baseline sensitivity loss (Fig. 3 depicts some examples). Therefore, when using Endpoint A, the required pre-selected visual field locations should be within −4.0 to -11.0 dB and within or adjacent to a scotoma present at baseline (i.e. study entry).

Discussion
In this retrospective, longitudinal study assessing real-world data of glaucomatous patients, we found that eyes experiencing rapid 24 - 2 visual field MD slopes over a 2-year period were more likely to reach one of the FDA accepted endpoints during or soon after that period. These findings suggest that enriching clinical trials with a sample of patients with progressing disease could enable the detection of clinically significant, regulatory compliant endpoints in glaucoma neuroprotection trials over a relatively short period of time.

The advantages of using MD slopes for reducing the sample size requirements in glaucoma trials have been recently discussed.\(^\text{15-19}\) Wu at al. have shown that the feasibility of glaucoma clinical trials could be improved by evaluating differences in the rate of visual field change (slopes) between randomization groups.\(^\text{19}\) For instance, assuming a between-group treatment effect of 30%, and a 90% statistical power, 1,924 participants would be required per group using more conventional event-based analysis, whereas 277 participants per group would be needed using trend analysis of the MD.\(^\text{19}\) Proudfoot et al. reported that for an 80% power to detect between-group differences in the rate of MD change could be attained with total follow-up between 18 months and 2 years and fewer than 300 total participants.\(^\text{18}\)

One main challenge facing neuroprotection clinical trials is the fact that all patients in such trials would be receiving some type of traditional IOP-lowering (i.e. standard care), which would increase the sample size and study duration requirements to detect a meaningful number of patients reaching a visual field endpoint.\(^\text{11}\) In the United Kingdom Glaucoma Treatment Study (UKGTS), for instance, about 23% of patients treated with a prostaglandin reached an event-based endpoint in 2 years.\(^\text{6}\) Nonetheless, that endpoint was a modification of the commercially available Glaucoma Progression Analysis (GPA) of the Humphrey Field Analyzer (HFA, Carl Zeiss Meditec, Inc., Dublin, CA, USA) and required three test locations to progress and be confirmed upon repeat testing. Both endpoints recommended by the FDA require at least 5 test locations, which would yield an even smaller number of eyes reaching an endpoint. Therefore, instead of an untreated placebo group (from an IOP standpoint), future trials need to consider enriching their population with patients likely to progress during the trial. There are many potential strategies for sample enrichment in glaucoma studies, such as using historical visual field or optical coherence tomography data, risk calculators, or excluding patients with low IOP and stable disease, among others. Regardless of the strategy chosen, which warrants further studies to assess feasibility and efficacy, the ultimate goal is to increase the number of patients reaching any given endpoint during the trial and thus increase the absolute effect size, which ultimately would reduce the sample size requirements. Our data suggests that if one such strategy enabled the enrollment of a sample progressing, on average, more rapidly than the mean MD slope of a general treated glaucoma population (-0.5 dB/year),\(^\text{7}\) around 70–80% of patients would reach an FDA recommended endpoint in 2 years, and almost 90% in 3 to 4 years. This translates into a significant reduction in sample size requirements. Additionally, the advantages of employing trend analysis of the visual field MD instead of pointwise event-based endpoints \(^\text{19}\) (discussed above) may further help reduce the costs of neuroprotection trials. Although a cut-off MD slope of -1.0 dB/year can help even further, the difference in predictive ability was minimal compared to -0.5 dB/year. This happened for two reasons: 1) there are far fewer patients progressing faster than -1.0 dB/year, and
2) all those patients progressing faster than −1.0 dB/year are also included in the group progressing faster than −0.5 dB/year. Finally, there are ethical issues in designing a trial where patients are losing visual field sensitivity at such rapid speed; they would likely require significant escalation of treatment with IOP lowering therapies during the trial, which could be a confounder in the analysis.

Some of the limitations of our study are its retrospective nature and the use of visual field data alone, without other clinical correlates. Nevertheless, this sample is drawn from a real-world population which is key for the translation to clinical practice. Despite the lack of other clinical data, all patients had or were suspected of having optic nerve damage and had confirmed glaucomatous visual field loss. In addition, masked review of the visual field data maintained data quality and mitigated the effect of non-glaucomatous effects on perimetric progression. As in previous clinical trials, future neuroprotection trials will also require masked reading centers ensuring the quality of the examinations and ruling out endpoints being met due to other causes. In the Ocular Hypertension Treatment Study (OHTS), for instance, it was shown that despite using an objective visual field endpoint, the role of a reading center was key to ensure the statistical power of the trial, without which the study would not have reached statistical significance. Finally, the frequency of testing in our sample (about two tests per year) is likely less than what a clinical trial should perform. Studies have shown that more frequent testing coupled with clustering at the beginning and end of the trial may further help improve the statistical power and reduce sample size requirements in glaucoma trials employing visual field tests and endpoint.

In summary, we found that neuroprotection trials employing visual field MD slopes could meet the FDA requirements for an alternative functional progression endpoint. The MD slopes had high predictability of both current and future progression based upon currently accepted endpoints. Such trials would require sample enrichment with patients progressing more rapidly than the average of the treated glaucoma population. This approach would allow for shorter duration trials with a significantly smaller sample size than those used in glaucoma trials looking at functional endpoints to date and help accelerate approval of novel therapies.

**Methods**

The Human Research Protection Office and Institutional Review Board at Columbia University approved the creation of the de-identified database of visual field tests included in this study. All methods were performed in accordance with the relevant guidelines and regulations of the Institutional Review Board. The study adheres to the tenets of the Declaration of Helsinki. Written informed consent was waived by the Human Research Protection Office and Institutional Review Board at Columbia University given the retrospective nature of the study.

Visual field tests were selected from the Edward S. Harkness Eye Institute at the Department of Ophthalmology at Columbia University Irving Medical Center. For this study, only the visual field data of patients with or suspected glaucoma (based on optic disc examination) were included. Only the visual
field data, without their clinical correlates, were available for this project. Patients in the dataset were treated at the discretion of attending physicians.

### Inclusion and Exclusion Criteria

Eyes with 24 − 2 visual field tests performed using the Swedish Interactive Thresholding Algorithm (SITA; Standard and Fast) with white-on-white stimuli of size III were selected for the present analyses. This initial dataset included 204,781 visual field tests from 34,362 patients with or suspected glaucoma based on the attending physicians’ assessment.

The following reliability criteria were applied to the visual field tests included in this study: less than 33% false negative responses, less than 20% false positive errors, and less than 33% fixation losses. To select patients with glaucomatous visual field loss, only eyes with a Glaucoma Hemifield Test (GHT) “Outside Normal Limits” or Pattern Standard Deviation (PSD) probability less than 0.05 on at least two consecutive visits were included. The first timepoint during the sequence of visual field tests when those criteria were met was defined as the baseline exam (Time = 0). All tests prior to the baseline exam were excluded from the analyses. Only eyes with baseline 24 − 2 MD between −3 and −15 dB were included. This range was chosen to minimize the chances of including eyes without established visual field loss as well as to mitigate the effects of severe functional loss on perimetric variability and the floor effect.

From the above, only eyes with at least seven reliable visual field tests were selected, of which the first five or more tests had to be done within 2 years. This sequence is herein called Sequence 1. This criterion was chosen to simulate a 2-year clinical trial while still taking into account that a minimum of five tests are needed for a reasonable calculation of slopes using linear regression. A minimum of two tests after that sequence would enable assessing changes from baseline in a follow-up examination and confirmation on a confirmatory test. This sequence after Sequence 1 is herein called Sequence 2. There were no constraints on the time between the first five visual field tests and the two or more after given the current recommendations do not impose a time limit on the predictability of either current or future endpoint assessment.

### Endpoint Analyses

To determine if the eye reached one of the recommended functional endpoints, the average of the first two tests (from Sequence 1) was used to define the pointwise baseline sensitivities. The total deviation (TD) values were used for these analyses because the current recommendations are based upon the GCP method (Carl Zeiss Meditec, Inc, Dublin, California, USA), which employs the TD instead of the pattern deviation (PD) plots for progression determination. Of note, the GCP algorithm is no longer commercially available, as it has been replaced by the Glaucoma Progression Analysis (GPA), which is based upon the PD plots. For consistency with the FDA's recommendations, we kept the analyses based upon the GCP in this study. The regression models to derive the GCP’s limits of test-retest variability are adapted from the work of Heijl et al.
For Endpoint A, that is, five or more locations worsening by at least 7 dB, if at any point in time the TD sensitivity decreased below that value and later confirmed in the next follow-up test, that test location was deemed progressing. If at least five locations met the above criteria on the same date, the eye was deemed progressing at that point in time. As per the agency’s recommendation, these five test locations need to be pre-defined at baseline. Given the challenges in predicting which test locations are going to deteriorate over time, we performed a post-hoc analysis identifying the characteristics of the locations more likely to progress based on that definition. This analysis included measures center (mean and median) and dispersion (interquartile range (IQR) and standard deviation (SD)). Note that, to meet the FDA’s recommendation of either current or future endpoint, the eye could reach the endpoint in either Sequence 1 (current) or 2 (future). If met in both, the earliest time point was chosen for these analyses.

In a similar fashion, for Endpoint B, that is, at least 5 test locations based upon the GCP algorithm, if at any point in time the TD sensitivity decreased below the GCP’s lower limit of test-retest variability and was confirmed in the next follow-up test, that test location was deemed progressing. If at least five locations met the above criteria on the same date, the eye was deemed progressing at that point in time. No constraints on the distribution of the progressing locations were required so as to resemble the GCP algorithm, that is, the five or more locations could be anywhere in the visual field.

Given the possibility that visual field changes due to non-glaucomatous causes could have detrimentally affected the determination of event-based progression (Endpoints An and B) as well as the calculation of MD slopes, a quality-check of all visual field sequences of the final dataset was performed by a glaucoma expert (CGDM) masked to the MD slopes and endpoint determination. If any non-glaucomatous changes in visual function were suspected (e.g., learning effect, fatigue, diffuse loss from cataract or retinal diseases), the eye was excluded from the final analyses.

**Statistical Analysis**

The visual field MD slopes of Sequence 1 were calculated using least squares linear regression, which is the statistical method most widely employed in commercially available machines. The coefficient ($\beta$) of the model: $MD \ (in \ dB) = \text{intercept} + \beta \times \text{Time \ (in \ years)}$ represents the slope (or speed of progression) in dB/year.

Kaplan-Meier survival models were used to describe the survival function of eyes reaching either endpoint criterion. The date of the first exam (of at least two) in which the eye reached one of the endpoints was used to define time to event. We investigated three MD slope cut-offs and their predictive value: (i) faster than 0 dB/year, i.e., describing eyes with declining sensitivity over time; (ii) faster than $-0.5$ dB, i.e., describing eyes progressing more rapidly than the average progression rate in treated glaucoma patients; and (iii) faster than $-1.0$ dB/year, i.e., describing eyes experiencing very rapid visual field progression. Cox Proportional Hazards models were employed to test for the statistical significance between these above MD slope cut-offs and the probability of reaching either endpoint, as well as to test for the magnitude of the associations. The length of time from the baseline exam until 25, 50, and 75% of eyes reached one of the endpoints was also assessed.
Computerized statistical analyses were performed with SAS version 9.4 (SAS Institute, Cary NC) and R (R Foundation for Statistical Computing, Vienna, Austria). Statistical significance was defined at P < 5%.

Abbreviations

intraocular pressure (IOP), mean deviation (MD), standard automated perimetry (SAP), Glaucoma Hemifield Test (GHT), Pattern Standard Deviation (PSD), Standard Deviation (SD), Ocular Hypertension Treatment Study (OHTS), United Kingdom Glaucoma Treatment Study (UKTGS).

Declarations

Conflict of Interest Statement: No conflicting relationship exists for any author.

Data availability

The datasets generated and/or analyzed during the current study are not publicly available due IRB guidance on data sharing but could be available from the corresponding author on reasonable request.

References


**Figures**

![Product-Limit Survival Estimates](image)

**Figure 1**

Survival curves depicting the survival probability of eyes reaching Endpoint A. Blue: all eyes; brown: eyes progressing faster than 0 dB/year; red: eyes progressing faster than -0.5 dB/year; and green: eyes progressing faster -1.0 dB/year. Time (x axis) in years.
Figure 2

Survival curves depicting the survival probability of eyes reaching Endpoint B. Blue: all eyes; brown: eyes progressing faster than 0 dB/year; red: eyes progressing faster than -0.5 dB/year; and green: eyes progressing faster -1.0 dB/year. Time (x axis) in years.

Figure 3

Location of progressing points as a function of baseline severity. The legend shows the relationship between baseline total deviation sensitivity (dB) and greyscale. The numbers inside each box (0 or 1) show whether that test location progressed based upon Endpoint 1. Left: both eyes of the same patient who reached the endpoint. Note that there are 5 locations adjacent to a deep baseline scotoma. Right: both eyes of the same patient who reached the endpoint. Note that there are 5 locations inside a relatively shallow baseline scotoma.