

Series Length used during Trend Analysis Affects Sensitivity to Changes in Progression Rate in the Ocular Hypertension Treatment Study (OHTS)

Stuart K. Gardiner, PhD *¹

Shaban Demirel, PhD ¹

Carlos Gustavo De Moraes, MD ^{2, 3}

Jeffrey M. Liebmann, MD ^{2, 3}

George A. Cioffi, MD ⁴

Robert Ritch, MD ^{2, 5}

Mae O. Gordon, PhD ^{6, 7}

Michael A. Kass, MD ⁶

for The Ocular Hypertension Treatment Study

¹ Devers Eye Institute, Legacy Health, Portland, OR.

² Einhorn Clinical Research Center, New York Eye & Ear Infirmary, New York, NY.

³ New York University School of Medicine, New York, NY.

⁴ Edward S. Harkness Eye Institute, Columbia University, New York, NY.

⁵ Department of Ophthalmology, The New York Medical College, Valhalla, NY

⁶ Ophthalmology & Vision Sciences, Washington University School of Medicine, St Louis, MO.

⁷ Division of Biostatistics, Washington University School of Medicine, St Louis, MO.

* Corresponding author:

Devers Eye Institute, Legacy Health,
1225 NE 2nd Ave, Portland, OR, 97232, USA.

Email: sgardiner@deverseye.org

Word Count: 3570

Abstract Word Count: 250

Figures: 2

Tables: 5

OHTS clinical trials identifier (www.clinicaltrials.gov): NCT00000125.

Financial disclosure:

None

Support:

NIH EY09307

NIH EY09341

National Center on Minority Health and Health Disparities

Merck, Inc., Whitehouse Station, NJ

Pfizer, Inc, New York, NY

Research to Prevent Blindness, New York, NY

Gildor Research Fund of the New York Glaucoma Research Institute, New York, NY

Glaucoma Research and Education Fund of Lenox Hill Hospital (CGDM), New York, NY

Legacy Good Samaritan Foundation, Portland, OR.

Abstract

Introduction

Trend analysis techniques to detect glaucomatous progression typically assume a constant rate of change. This study uses data from the Ocular Hypertension Treatment Study to assess whether this assumption decreases sensitivity to changes in progression rate, by including earlier periods of stability.

Methods

Series of visual fields (mean 24 per eye) completed at 6-month intervals from participants initially randomized to observation were split into subseries before and after the initiation of treatment (the “split-point”). The Mean Deviation Rate of change (MDR) was derived using these entire subseries, and using only the W tests nearest the split-point, for different window lengths of W tests. A generalized estimating equation model was used to detect changes in MDR occurring at the split-point.

Results

Using shortened subseries with $W=7$ tests, the MDR slowed by 0.142dB/yr upon initiation of treatment ($p<0.001$), and the proportion of eyes showing “rapid deterioration” (MDR<-0.5dB/yr with $p<5\%$) decreased from 11.8% to 6.5% ($p<0.001$). Using the entire sequence, no significant change in MDR was detected ($p=0.796$), and there was no change in the proportion of eyes progressing ($p=0.084$). Window lengths $6\leq W\leq 9$ produced similar benefits.

Conclusion

Event analysis revealed a beneficial treatment effect in this dataset. This effect was not detected by linear trend analysis applied to entire series, but was detected when using shorter subseries of length between six and nine fields. Using linear trend analysis on the entire field sequence may not be optimal for detecting and monitoring progression. Non-linear analyses may be needed for long series of fields.

Introduction

The Ocular Hypertension Treatment Study (OHTS) was a randomized clinical trial that demonstrated the beneficial effect of intraocular pressure (IOP) reduction to prevent or delay the onset of primary open-angle glaucoma.^{1, 2} A second phase of the trial in which all participants were offered IOP-lowering medication, demonstrated that the rate of conversion to primary open-angle glaucoma (POAG) was lower in those participants who had been initially randomized to treatment, i.e. that delaying treatment in the observation group negatively affected the average status, especially in higher-risk participants.³

Major clinical trials in glaucoma, including OHTS, have formulated different criteria for classifying cross-sectional disease status and have used change in status as a basis for detecting change over time.^{1, 4-7} However, such event-based criteria do not make full use of the available information; partly because they only produce a binary outcome (in this case, *reaching POAG endpoint or not*) and partly because they do not use all tests carried out during the follow-up series. Additionally, some forms of event analysis define progression as a change in status from within to outside normal limits. These methods are less sensitive for participants whose baseline status is toward the upper end of the normal range. Therefore, trend analysis for evaluation of both structural and functional damage in glaucoma has become the subject of increasing interest among clinicians and researchers.^{8, 9} Rates of change provide objective, continuous variables quantifying progression, making use of all available data in the series. Such extra information about the rate at which patients progress may help predict future functional loss and vision-related quality of life, and aid clinicians when assigning management strategies for their patients. We have previously used trend analysis to show that participants reaching an event-analysis-based POAG endpoint in OHTS had significantly more rapid rates of functional change than those who did not,¹⁰ indicating that the rate of change is an effective measure of progression. We have also shown that treatment significantly slowed the rate of progression.¹¹

In light of the increasing importance of trend-based analysis methods in glaucoma, it is imperative that those methods be validated and optimized as much as possible. OHTS provides an invaluable resource in this regard, since it contains bi-annual visual field (VF) test results for over one thousand participants with a median follow-up of thirteen years.³ An effective trend-based analysis method should be able to detect changes in the rate of change while also having good specificity. The main OHTS outcome confirmed that initiation of ocular hypertension treatment improved outcomes (decreasing the probability of subsequently reaching a POAG endpoint),² and so an effective analysis should be able to detect a slowing in rate of change occurring at this time point. For participants initially randomized to the observation group, the rate of decline in VF status should slow after the initiation of treatment.

Participants had normal VF tests at enrollment into OHTS. Consequently, an initial sequence of normal fields could reduce the magnitude of the rate of change of sensitivity prior to initiation of treatment. This could reduce the ability to detect a change in rate occurring after treatment commenced. In this study, we examine the effect of series length on the ability to detect a change in rate occurring at the time of initiation of treatment, when using linear trend analysis techniques.

Methods

Baseline data and design of the OHTS have been described elsewhere.^{1, 2} All OHTS participants signed a statement of informed consent prior to study entry after having the risks and benefits of participation explained to them. The institutional review boards at each participating clinical site approved their respective informed consent statements and procedures. The study adhered to the tenets of the Declaration of Helsinki.

All subjects enrolled in the OHTS had to have at least two reliable (fixation losses, false-negative, and false-positive responses <33%), achromatic, automated VF test results (Humphrey Field Analyzer using the 30-2 testing pattern, Carl Zeiss Meditec, Inc., Dublin, CA) that were within normal limits during the qualifying period. The OHTS analysis dataset available for this study contained all VF tests and endpoint determinations in the OHTS database as of March 2009. From the full OHTS dataset ($n = 1636$), we first removed from further consideration any eye that reached an endpoint that was determined by the endpoint committee to be due to causes other than POAG (261 eyes of 202 subjects). We then selected only those follow-up VF tests that were considered reliable (false positives, false negatives and fixation loss all <33% if the Full Threshold Algorithm was used, false positives <15%, false negatives and fixation loss <33% if the Swedish Interactive Threshold Algorithm, SITA, was used).

A “Delayed Treatment Cohort” was formed, consisting of those eyes that had a change from observation to treatment at some point during follow-up. This analysis included eyes initially randomized to observation and which began treatment either as a result of a POAG endpoint determination or at the transition between the first and second phases of OHTS (at which time treatment was offered to all originally untreated subjects in the observation arm). The visit at which an eye was first noted to be on treatment became known as the “split-point” for that eye.

For a given window length of W tests, four sequences of VF tests were analyzed per subject:

- Before _{W} : the W most recent VF tests prior to the split-point
- After _{W} : the first W VF tests beginning at least 9 months after the split-point
- Before_{All}: all VF tests prior to the split-point
- After_{All}: all VF tests beginning at least 9 months after the split-point

A 9-month gap was left between commencing treatment and the start of the “after” sequence to ensure that the participant’s treatment had a chance to stabilize, allowing their physician to

determine a drug and dosage that resulted in attaining the target IOP. Subjects with fewer than W VF tests before and W VF tests after the split-point were excluded. We did not require that treatment be continuous once commenced.

Secondly, a “Continuous Treatment Cohort” was formed consisting of those eyes having no change in treatment status during the study; this therefore consisted of subjects initially randomized to treatment. For this cohort, the series was split chronologically into two equal parts, with a “split-point” halfway through the series. In this case, no gap to allow treatment to stabilize is needed. Therefore, the sequence $After_{All}$ consisted of all VF tests after the split-point.

Linear regression of mean deviation (MD) over time was performed separately for each of the four sequences for each of the eyes selected. MDs were considered to be equivalent between the two testing algorithms. The rate of change of MD [MDR (dB/yr)] was recorded for each eye, together with the standard error of the slope estimate, and a determination of whether “rapid deterioration” occurred during that period (defined as an MDR worse than -0.5 dB/yr that was significantly negative with $p < 0.05$).

An effect of the initiation of treatment on MDR was sought using data from both eyes of each individual (where available). A paired comparison was performed to determine the change in MDR that occurred at the split-point. Specifically, the change in MDR given by MDR “after” minus MDR “before” was set as the outcome of a generalized estimating equation (GEE) regression¹² with no independent variables. The resulting intercept term and p-value were used as estimates of the average change in MDR and its level of significance, respectively. This is analogous to a paired t-test comparing the two MDRs, but using a GEE regression to account for the fact that there may be correlated data from two eyes for the same participant. Additionally, the proportion of eyes for which

“rapid deterioration” occurred was calculated. These proportions were compared between the “Before” and “After” series using McNemar’s test.

This analysis was repeated for different window lengths W , ranging from $W=4$ (assumed to be the shortest window length over which linear regression can provide a reasonable estimate of the rate of change) to $W=12$ (the longest window length for which there were sufficient eyes with enough VF tests both before and after their split-point). Note that as the window length W increases, the number of eyes for which there are sufficient VF tests in both sequences is reduced; therefore the average MDR in the $\text{Before}_{\text{All}}$ and $\text{After}_{\text{All}}$ sequences varies depending on the value of W .

Results

Table 1 gives the change in MDR at the split-point for the Delayed Treatment Cohort, together with the significance level of that change, both for the shortened and complete sequences, for different window lengths W . In each case, only eyes with at least W VF tests in both their sequences Before_W and After_W were included. This means that the number of eligible eyes (n) and the average change in MDR observed in the entire sequence of fields both vary with W , as seen in the Table, due to different subsets of the complete dataset being eligible. It can be seen that using the shortened sequence length, a significant change in MDR is observed at the split-point, with the MDR in sequence After_W being less rapid than the MDR in sequence Before_W . This is consistent with the main finding of OHTS that a significant and beneficial treatment effect occurs. However, when the entire series of VF tests is used, there is no significant change in MDR at the split-point between sequences $\text{Before}_{\text{All}}$ and $\text{After}_{\text{All}}$. As W increases, the magnitude of the apparent improvement in MDR decreases, until eventually it becomes non-significant when $W>10$.

<i>W</i>	<i>n</i>	Shortened Sequence			Entire Sequence		
		Before _{<i>W</i>}	After _{<i>W</i>}	Significance	Before _{All}	After _{All}	Significance
4	892	-0.230 dB/yr	-0.055 dB/yr	0.001	-0.084 dB/yr	-0.118 dB/yr	0.203
5	884	-0.265 dB/yr	-0.099 dB/yr	<0.001	-0.083 dB/yr	-0.119 dB/yr	0.189
6	858	-0.249 dB/yr	-0.094 dB/yr	<0.001	-0.084 dB/yr	-0.107 dB/yr	0.275
7	802	-0.231 dB/yr	-0.089 dB/yr	<0.001	-0.087 dB/yr	-0.091 dB/yr	0.796
8	720	-0.204 dB/yr	-0.075 dB/yr	<0.001	-0.090 dB/yr	-0.085 dB/yr	0.812
9	601	-0.192 dB/yr	-0.063 dB/yr	<0.001	-0.096 dB/yr	-0.075 dB/yr	0.289
10	394	-0.196 dB/yr	-0.077 dB/yr	<0.001	-0.102 dB/yr	-0.088 dB/yr	0.557
11	157	-0.172 dB/yr	-0.110 dB/yr	0.206	-0.120 dB/yr	-0.119 dB/yr	0.978
12	73	-0.181 dB/yr	-0.190 dB/yr	0.905	-0.167 dB/yr	-0.197 dB/yr	0.655

Table 1: the mean MDR (slope of Mean Deviation over time) in the sequence of fields before the split-point and in the sequence of fields starting at least 9 months after the split-point for the Delayed Treatment Cohort using different window lengths *W*. Results are presented both when using only the *W* fields before and the *W* fields after the split-point for each sequence (“Shortened Sequence”), and when using all available fields (“Entire Sequence”). Note that the number of eyes included in the analysis (*n*) varies with *W*, since there must be at least *W* fields in both the Before and After sequences for an eye to be included; therefore the results using the entire sequence also vary with *W*. “Significance” gives the p-value indicating whether the change in MDR is significantly different from zero, using a generalized estimating equation model. Testing was carried out every six months; therefore the time interval covered by the shortened sequences varied from 1.5 years (*W*=4) to 5.5 years (*W*=12).

Table 2 shows the standard error of the estimates of MDR, averaged over all eyes and averaged over the two sequences (before and after the split-point). It is seen that when *W* is small, the error about the estimate of MDR in the shortened sequences is large. As the sequence length increases, the MDR can be estimated more accurately (the standard error decreases). Figure 1 shows box-and-whisker plots of the change in MDR for the shortened sequences, for all eyes, showing the variability decreasing as *W* increases.

W	Shortened Sequence	Entire Sequence
4	0.634 dB/yr	0.161 dB/yr
5	0.463 dB/yr	0.160 dB/yr
6	0.364 dB/yr	0.156 dB/yr
7	0.296 dB/yr	0.150 dB/yr
8	0.250 dB/yr	0.148 dB/yr
9	0.212 dB/yr	0.141 dB/yr
10	0.186 dB/yr	0.136 dB/yr
11	0.177 dB/yr	0.141 dB/yr
12	0.179 dB/yr	0.154 dB/yr

Table 2: The standard error of the MDR estimates (the slope from a linear regression of Mean Deviation over time), averaged over all sequences, when different window lengths W were used. The sample size n varies with W , as reported in Table 1.

Combining these two effects, Table 3 shows the proportion of eyes in the Delayed Treatment Cohort showing “rapid deterioration” before and after the split-point. It is seen that when W is too small, the variability in the slope estimate (as seen in Table 2) is so high that relatively few slopes reach a significance level of $p < 5\%$, and so fewer eyes are detected as showing “rapid deterioration”. When W is too large, the MDR is less likely to reach -0.5dB/yr over that time window, due to changes in the progression rate over that period, and so eyes that progress intermittently or for shorter periods are again missed. There is then no significant difference between the proportions of eyes flagged for “rapid deterioration” before and after the split-point, despite the presence of a known treatment effect that should be reducing the proportion of progressing eyes. When W is between approximately 6 and 9 (equivalent to 2.5 years and 4 years of follow-up respectively), the proportion of eyes detected as showing “rapid deterioration” before treatment was initiated is more than double the number detected when using the entire sequence. The proportion of eyes showing “rapid deterioration” in sequence Before_W was significantly higher than the proportion in sequence

Before_{All} for window lengths $5 \leq W \leq 9$ ($p < 0.01$ in each case). When $W=10$, $p=0.01$; when $W=11$, $p=0.15$; and when $W=12$, $p=1.00$. The proportion showing “rapid deterioration” in sequence After_W was lower than in sequence After_{All} when $W=4$ ($p < 0.01$) but not significantly different for other window lengths.

W	Shortened Sequence			Entire Sequence		
	Before _W	After _W	p	Before _{All}	After _{All}	p
4	5.0%	3.1%	0.054	4.7%	6.8%	0.065
5	7.9%	5.1%	0.019	4.6%	6.7%	0.076
6	10.6%	6.4%	0.002	4.8%	6.8%	0.093
7	11.8%	6.5%	<0.001	4.6%	6.7%	0.079
8	10.4%	7.2%	0.040	4.3%	6.4%	0.101
9	11.1%	5.8%	0.002	4.7%	6.2%	0.306
10	9.6%	5.6%	0.053	4.8%	6.1%	0.522
11	10.8%	8.3%	0.540	5.7%	8.3%	0.480
12	9.6%	8.2%	1.000	9.6%	9.6%	1.000

Table 3: The proportion of eyes showing “rapid deterioration” (MDR worse than -0.5dB/yr , significantly negative with $p < 5\%$) in the Delayed Treatment Cohort, together with the p-value comparing the proportions before and after the split-point (using McNemar’s test), both for shortened sequences of W fields before and after the split-point and when using the entire sequences. As before, results for the entire sequence vary slightly with W due to differences in the number of sequences eligible for analysis (n in Table 1).

As noted in the Methods section, the number of available series depends on the window length W . To ensure that this did not bias the results, the analysis was repeated using different lengths W but in each case only using the 394 series for which there were at least ten examinations both before and after the split-point. This minimum of 10 fields before the split-point also removes series for which treatment was initiated following an endpoint. As seen in Table 1, using $W=10$ reveals a

change in average MDR of 0.118dB/yr at the split-point. Consistent with the results above, using shorter windows revealed greater improvements in MDR, of 0.171dB/yr for $W=8$, 0.272dB/yr for $W=6$, and 0.365dB/yr for $W=4$.

Figure 2 shows the series of MD for a sample participant in the study. A relatively stable period over the first few years of the study was followed by a period of rapid progression, resulting in POAG endpoint determination (by both visual field and optic disc) and subsequent initiation of treatment at the date indicated by the gray vertical line. Using the entire sequence, the MDRs (the slopes of the blue lines in Figure 2) were -0.540dB/yr before the split-point and -0.542dB/yr over the fields at least nine months after the split-point. Using the shortened sequence of length $W=7$ fields, the MDRs (the slopes of the red lines in Figure 2) were -1.187dB/yr in the last seven fields before the split-point and -0.558dB/yr in the first seven fields after. The initial stable period and the later period of more rapid deterioration are combined, causing the overall rate to be less severe when using the entire sequence, making trend analysis less sensitive to the rapid progression in the later part of the pre-treatment period.

Table 4 shows the change in MDR occurring at the split-point for the Continuous Treatment Cohort, in the same format as Table 1. These participants were offered treatment throughout both phases of the trial (hence they have relatively slow rates of change). It is seen that using the shortened sequence, no significant change in rate is detected at the split-point. There is no reason to expect there to be a change in MDR when the split point is not associated with a change in treatment.

Table 5 shows, in the same format as Table 3, the proportion of eyes in the Continuous Treatment Cohort showing “rapid deterioration.” Again, no significant change in the proportion of eyes with “rapid deterioration” is detected using the shorter sequences. Note that just because these participants received treatment throughout the study does not guarantee that they had no glaucomatous progression. Out of the 1970 eyes in the Continuous Treatment Cohort (not all of

which had at least 4 VFs for both their Before_W and After_W series), 13 (1.4%) reached a glaucomatous endpoint.

W	n	Shortened Sequence			Entire Sequence		
		Before _W	After _W	Significance	Before _{All}	After _{All}	Significance
4	1921	-0.082 dB/yr	-0.069 dB/yr	0.731	-0.039 dB/yr	-0.125 dB/yr	<0.001
5	1890	-0.077 dB/yr	-0.063 dB/yr	0.631	-0.036 dB/yr	-0.119 dB/yr	<0.001
6	1797	-0.081 dB/yr	-0.058 dB/yr	0.376	-0.040 dB/yr	-0.113 dB/yr	<0.001
7	1730	-0.071 dB/yr	-0.057 dB/yr	0.541	-0.040 dB/yr	-0.107 dB/yr	<0.001
8	1617	-0.063 dB/yr	-0.050 dB/yr	0.540	-0.040 dB/yr	-0.093 dB/yr	0.004
9	1508	-0.064 dB/yr	-0.062 dB/yr	0.919	-0.041 dB/yr	-0.087 dB/yr	0.010
10	1390	-0.042 dB/yr	-0.058 dB/yr	0.368	-0.035 dB/yr	-0.081 dB/yr	0.005
11	1231	-0.022 dB/yr	-0.054 dB/yr	0.034	-0.024 dB/yr	-0.072 dB/yr	<0.001
12	986	-0.014 dB/yr	-0.056 dB/yr	0.005	-0.017 dB/yr	-0.068 dB/yr	<0.001

Table 4: The mean MDR (slope of Mean Deviation over time) in the sequence of fields before the split-point and in the sequence of fields after the split-point for the Continuous Treatment Cohort, using different window lengths *W*. Results are presented both when using only the *W* fields before and the *W* fields after the split-point for each sequence (“Shortened Sequence”), and when using all available fields (“Entire Sequence”). “n” shows the number of eyes included in the analysis, which varies with *W* (as in Table 1). “Significance” gives the p-value indicating whether the change in MDR is significantly different from zero, using a generalized estimating equation model.

<i>W</i>	Shortened Sequence			Entire Sequence		
	Before _{<i>W</i>}	After _{<i>W</i>}	p	Before _{All}	After _{All}	p
4	3.7%	3.1%	0.287	4.3%	7.9%	<0.001
5	4.0%	4.3%	0.687	4.3%	8.0%	<0.001
6	5.3%	4.9%	0.644	4.4%	7.9%	<0.001
7	5.9%	6.0%	0.942	4.5%	7.5%	<0.001
8	5.9%	6.4%	0.608	4.3%	6.7%	0.002
9	5.4%	6.3%	0.310	3.9%	6.6%	0.001
10	4.5%	5.3%	0.327	3.6%	6.2%	0.001
11	3.5%	4.5%	0.198	3.3%	5.4%	0.010
12	3.5%	4.5%	0.336	2.9%	5.4%	0.006

Table 5: The proportion of eyes showing “rapid deterioration” in the Continuous Treatment Cohort (MDR worse than – 0.5dB/yr, significantly negative with $p < 5\%$), together with the p-value comparing the proportions before and after the split-point (using McNemar’s test), both for shortened sequences of *W* fields before and after the split-point and when using the entire sequences.

Discussion

Our results indicate that when linear trend analysis techniques are used, using the entire available sequence of visual fields can decrease the sensitivity to detecting known changes in the rate of functional progression. Shorter sequences of only the more recent fields prior to the split-point make linear trend analysis more sensitive to such changes, without compromising specificity. Series of between six and nine VFs provided the best sensitivity in this analysis. While shorter sequences could be expected to result in more “false positive” cases wherein progression is flagged in stable eyes, no evidence of a significant reduction in specificity was found. To our knowledge, this is the first report suggesting that using the entire VF sequence may not be optimal for detecting and monitoring progression when using linear trend analysis techniques.

An implication of these findings is that the linear model for progression is suboptimal for long series of visual fields in ocular hypertensive eyes. This is not surprising, since such a model assumes a constant rate of change. Consider a 60-year old patient, currently with an MD of -5dB, progressing at -1.0dB/yr. Using a linear model and extrapolating the trend would imply that at the age of 40 their MD had been +15dB, which is clearly nonsensical. Similarly, if a patient's disease has progressed to the extent that their sensitivity has reached 0dB, no further decline corresponding to a linear model is possible. A linear model for progression would overestimate the amount of change in sensitivity once this floor has been reached. Non-linear models for progression that allow for periods of zero apparent change both at the start and at the end of a series are needed. The OHTS dataset used for this study does not contain a full range of disease severities (since very few subjects experienced severe vision loss by the end of the study), and so no specific non-linear model can be advocated at this stage without the necessary data for validation. In the meantime, if linear models for progression are to be used, it is recommended that the sequence be plotted against time and visually inspected to detect acceleration (for individual patients), and/or analysis restricted to the more recent visual fields in the series (especially in the case of clinical studies). In addition, the use of a linear model to fit data measured on a logarithmic scale (dB) should be interpreted with caution. Although it simplifies how clinicians measure and understand progressive visual field change, one should be reminded that a given amount of dB loss in the early stages of the disease (e.g.: 1.0 dB) corresponds to a smaller absolute sensitivity change than in later stages for the same dB loss. Using shorter series of tests should minimize this confounding effect when measuring rates of change in dB/yr in eyes with normal to early damage in the visual field. This hypothesis needs to be tested in eyes with advanced field loss.

Clinical assessment of progression in glaucoma has included seeking evidence of acceleration in the rate of progression. Even if perimetry were free of variability, so that the rate of change could be known precisely, it still could not be assumed that a patient's rate of change is constant, with no

need for future testing. A previously stable patient can start to progress rapidly, and we cannot yet predict the onset of this change. Therefore in some respects, our findings are more in agreement with current clinical practice than the principle of using a patient's entire series of fields. Indeed, the EyeSuite program developed to analyze longitudinal series of results from the Octopus perimeter (Haag Streit International, Berne, Switzerland) by default assesses the rate of change over the most recent six tests, rather than the entire series. While it may seem counter-intuitive to not use all the data (especially in light of the fact that the biggest advantage of trend analysis methods is that they make better use of the available data than event analysis techniques), as seen in the example in Figure 2 there is a sound logical underpinning behind assessing the relatively recent rate of change.

The optimum series length to maximize sensitivity when aiming to detect progression will depend on many factors. In this study, series of six to nine VFs provided good sensitivity when tests were carried out at six-month intervals, spanning a period of 2.5 to 4 years. If testing were carried out only annually (as is often the case for ocular hypertensive patients), four years may be insufficient, since it would provide a series of only five fields. However, the optimum may still be to analyze subseries of fewer than nine fields in order to detect shorter-term changes in the progression rate. By contrast, if three tests were to be performed annually, as has been recommended for some purposes,⁹ it may be possible to increase the series length (hence reducing variability about the estimate of rate of change) without compromising sensitivity of detecting rapid progression. Another factor to be taken into consideration is that some patients produce more variable VFs than others, and so may require longer series in order for progression to become apparent.

The optimum analysis method and optimum series length could depend on disease severity. The main justification for not always using the first few fields in the series with linear analysis is that the patient may be stable for some time before progression begins, as in the example in Figure 2. A patient who has already developed a glaucomatous defect would be considered less likely to have a

prolonged period of stability before progression accelerates. In addition, variability is much higher in more advanced disease, potentially making estimates of the rate of change based on fewer fields unreliable. As variability increases, robustness becomes more important, favoring linear models over non-linear models with higher numbers of free parameters, but using longer series of fields to obtain more accurate estimates of the rate of change. Since this dataset does not contain a large number of cases of moderate or severe glaucoma, this conjecture would need testing in a different dataset.

Using the entire sequence, a significant change in MDR was observed at the split-point in the Continuous Treatment Cohort, with these eyes progressing more rapidly in the second half than in the first half of the study. No change in treatment status occurred during their sequence, as all patients in this cohort were treated from the start of the study (although the treatment given may have changed). It is possible that this is a chance characteristic of the data. However, even though the sample size is reduced when W is large, 986 eyes would still be considered more than adequate. It may be that this effect is caused by a significant number of those eyes beginning to progress towards the end of the sequence. It could also be indicative of non-linearity of progression, with sensitivities accelerating downwards, as would be consistent with our previous findings in another dataset that the current MD is predictive of the rate of subsequent change.¹³ Finally, it would also be consistent with the presence of a learning effect causing sensitivities to rise over the first few fields of the series;¹⁴ however, all OHTS participants were required to have had previous fields indicating at least some familiarity with automated perimetry. Although we refer to this group as the Continuous Treatment Cohort because their treatment status did not change, it cannot be assumed that they were not undergoing any glaucomatous progression. The fact that progression may be accelerating in many of the eyes in the Continuous Treatment Cohort that have been consistently managed, biasing the results towards a more rapid MDR later in the study, makes it all

the more impressive that initiating treatment had such a clear beneficial effect in the Delayed Treatment Cohort.

A caveat with the findings is that commencement of treatment could have been due to the participant reaching an endpoint in the first stage of the study, or could have been due to the decision to offer treatment to everyone in the second stage of the study. Participants who reached an endpoint in the first phase will have shorter sequences $Before_W$ available, and so would not be eligible for inclusion when using longer window lengths W . This could partly explain the greater change in MDR when using, for example, $W=4$ instead of $W=8$. However, as W becomes quite large (7 or 8 fields), such participants will form only a small proportion of the sample size n , and so it is unlikely that this is driving the main conclusions of the analysis. Notably, when the same analysis was performed varying W but consistently using all series with at least 10 fields before and after the split-point, the change in MDR was still greater for smaller W .

This study used MD to generate a measure of the rate of functional change, corresponding to disease progression. MD is useful as a global measure in clinical trials such as the OHTS, but is insensitive to deterioration of small scotomas in individual patients. Clinically, change in MD would be just one of several measures used to determine whether changes in treatment are necessary. Pointwise changes are more variable. However, the same principle would apply, and non-linear methods developed for MD are likely to have similar benefits when applied to pointwise data.

The main conclusion to be drawn from this study is that using the entire series of test results for linear trend analysis may actually be detrimental to early detection of rapid visual field change, especially when that progression is sporadic or preceded by a period of stability. At this early disease stage, rates of change were consistently underestimated when the entire sequence was used. Use of shorter sequences improved the ability to detect slowing of the rate of progression at

the time treatment was initiated. By contrast, use of the shorter sequences did not cause a significant increase in the number of series for which a change in rate was detected in the absence of a change in treatment status. These results underscore the need for non-linear models for progression, while also providing a method to reduce the problem until such models have been developed and validated. Such techniques could make trend analysis more sensitive to changes in the rate of progression, allowing earlier detection and implementation of appropriate treatments.

References

1. Gordon MO, Kass MA, for the Ocular Hypertension Treatment Study G. The Ocular Hypertension Treatment Study: Design and Baseline Description of the Participants. *Arch Ophthalmol* 1999;117:573-583.
2. Kass M, Heuer D, Higginbotham E, et al. 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.
3. Kass MA, Gordon MO, Gao F, et al. Delaying Treatment of Ocular Hypertension: The Ocular Hypertension Treatment Study. *Arch Ophthalmol* 2010;128:276-287.
4. The Advanced Glaucoma Intervention Study Group. The Advanced Glaucoma Intervention Study (AGIS): 1. Study design and methods and baseline characteristics of study patients. *Controlled Clinical Trials* 1994;15:299-325.
5. Leske M, Heijl A, Hyman L, Bengtsson B. Early manifest glaucoma trial - Design and baseline data. *Ophthalmology* 1999;106:2144-2153.
6. Musch D, Lichter P, Guire K, Standardi C, Group TCS. The Collaborative Initial Glaucoma Treatment Study: study designs, methods, and baseline characteristics of enrolled patients. *Ophthalmology* 1999;106:653-662.
7. Miglior S, Zeyen T, Pfeiffer N, Cunha-Vaz J, Torri V, Adamsons I. The European glaucoma prevention study design and baseline description of the participants. *Ophthalmology* 2002;109:1612-1621.

8. Caprioli J. The Importance of Rates in Glaucoma. *American Journal of Ophthalmology* 2008;145:191-192.
9. Chauhan B, Garway-Heath D, Goni F, et al. Practical Recommendations for Measuring Rates of Visual Field Change in Glaucoma. *Br J Ophthalmol* 2008;92:569-573.
10. Demirel S, De Moraes CGV, Gardiner SK, et al. The Rate of Visual Field Change in the Ocular Hypertension Treatment Study. *Invest Ophthalmol Vis Sci* 2012;53:224-227.
11. De Moraes CG, Demirel S, Gardiner SK, et al. Effect of Treatment on the Rate of Visual Field Change in the Ocular Hypertension Treatment Study Observation Group. *Invest Ophthalmol Vis Sci* 2012;53:1704-1709.
12. Liang K, Zeger S. Longitudinal data analysis using generalized linear models. *Biometrika* 1986;73:13-22.
13. Gardiner S, Demirel S, Johnson C. Perimetric indices as predictors of future glaucomatous functional change. *Optom Vis Sci* 2011;88:56-62.
14. Gardiner S, Demirel S, Johnson C. Is There Evidence for Continued Learning Over Multiple Years in Perimetry? *Optom Vis Sci* 2008;85:1043-1048.

Figure Legends

Figure 1: Box-and-whisker plots of the change in mean deviation rate (MDR) occurring at the split-point, for different lengths of shortened window W . For each box, the central horizontal line represents the median value; the box covers the inter-quartile range; and the whiskers extend to the maximum and minimum values. The gray horizontal line indicates zero difference in MDR.

Figure 2: The series of visual field test results for a sample participant in the Delayed Treatment Cohort. The vertical gray line represents the first visit at which the subject was receiving treatment. The blue lines show the rate of change of mean deviation from linear regression over the entire sequences, before the split-point and commencing at least 9 months after the split-point. The red lines show the equivalent rates of change derived using the shortened sequences, i.e. the last seven fields before the split-point and the first seven fields at least 9 months after the split-point. Before the split-point, the red line is steeper than the blue line, showing that use of the shortened sequence makes trend analysis more sensitive to the rapid rate of change that occurred over this period. A change in rate is apparent at treatment onset using the shortened sequence, but not using the entire sequence. Note that when the split-point was caused by the subject reaching an endpoint, all confirmation fields were included in the Before _{W} sequence.

Change in MDR at split-point (dB/yr)



