Intervals between Visual Field Tests When Monitoring the Glaucomatous Patient: Wait-and-See Approach

David P. Crabb[1] *and David F. Garway-Heath*[2]

Purpose. Published recommendations suggest three visual field (VF) tests per year are required to identify rapid progression in a newly diagnosed glaucomatous patient over 2 years. This report aims to determine if identification of progression would be improved by clustering tests at the beginning and end of the 2-year period.

Methods. Computer-simulated "patients" were given a rapid VF (mean deviation [MD]) loss of −2 dB/year with added MD measurement variability. Linear regression of MD against time was used to estimate progression. One group of "patients" was measured every 6 months, another every 4 months, whereas the wait-and-see group were measured either 2 or 3 times at both baseline and at the end of a 2-year period. Stable "patients" (0 dB/year) were generated to examine the effect of the follow-up patterns on false-positive (FP) progression identification.

Results. By 2 years, 58% and 82% of rapidly progressing patients were correctly detected using evenly spaced 6- and 4-month VFs, respectively. This power of detection significantly improved to 62% and 95% with the wait-and-see approach ($P < 0.001$). When compared with evenly spaced VFs, the rate of MD loss was better estimated by the wait-and-see approach, but average detection time was slightly slower. Evenly spaced testing incurred a significantly higher FP rate: up to 5.9% compared with only 0.4% in wait-and-see ($P < 0.001$).

Conclusions. Compared with an evenly spaced follow-up, wait-and-see identifies more "patients" with rapid VF progression with fewer FPs, making it particularly applicable to clinical trials. Modeling experiments, as reported here, are useful for investigating and optimizing follow-up schemes. (*Invest Ophthalmol Vis Sci.* 2012;53:2770–2776) DOI:10.1167/iovs.12-9476

Accurately determining visual field (VF) progression in glaucoma is important: it is central to effective clinical management of the disease, and it is the relevant functional endpoint for clinical trials of new treatments. Monitoring VF status in the patient is done by testing over a period of time. Appropriate analysis to measure VF progression then requires a sound comprehension of the amount or rate of VF loss (decibels [dB] per year), the period of observation, the effect of VF measurement variability, and the number of follow-up tests required to detect change with adequate statistical power.[1] The last of these is often overlooked. VFs can exhibit extensive measurement error and it is widely accepted that clinical management decisions made on results from one or two tests will as a rule be unsound. A sufficient observation period and an adequate number of VFs are required before change can be documented with confidence. Furthermore, longitudinal studies[2–4] and snippets of natural history data[5,6] indicate there is significant between-patient variability in the rate of VF progression. Rates of VF loss in individuals cannot be well predicted with knowledge about the patient at diagnosis, but might be best estimated by collecting VFs over a period of time. In a newly diagnosed patient the ideal would be to obtain an adequate number of reliable examinations over a period of time to exclude, or to detect, the presence of rapid VF progression, thus offering the chance of intensifying treatment to control the disease process in those patients that need it. Until recently there has been little research evidence to offer guidance concerning how frequently VF tests should be done to optimally detect progression or best estimate the rate of loss. There is certainly a need for an evidence base for this issue because the scheduling of VFs in routine clinical practice is often erratic and not well planned.[7]

Chauhan and colleagues (2008) published useful practical recommendations for measuring rates of VF change in glaucoma based on statistical power calculations.[8] Central to their report was the important principle that a sufficient number of tests must be performed in a specified period to give any chance of separating true disease progression from the measurement variability inherent in the VF data. This sensible conclusion is analogous to the accepted idea that a clinical trial will not be sufficiently powered to detect an experimental effect if an insufficient number of patients is recruited. One specific result from the report suggested that newly diagnosed patients should be tested with standard automated perimetry (SAP) three times per year during the first 2 years after diagnosis. By using this approach, rapidly progressing eyes, losing average VF sensitivity (mean deviation [MD]) of more than 2 dB per year, will be indentified with greater certainty than if less frequent testing is performed. These recommendations have subsequently been adopted as guidelines by the European Glaucoma Society in their advice on patient

From the [1]Department of Optometry and Visual Science, City University London, London, United Kingdom; and the [2]National Institute for Health Research Biomedical Research Centre for Ophthalmology, The Moorfields Eye Hospital/National Health Service Foundation Trust and University College London Institute of Ophthalmology, London, United Kingdom.

This work was presented in part at the Association of Research in Vision and Ophthalmology Meeting, Fort Lauderdale, Florida, May 2009.

Supported in part by the UK National Institute for Health Research (NIHR) Health Services Research programme (project 10/2000/68), the International Glaucoma Association and the Department of Health through the award made by the NIHR to Moorfields Eye Hospital NHS Foundation Trust and UCL Institute of Ophthalmology for a Biomedical Research Centre for Ophthalmology. The views expressed in this publication are those of the authors and not necessarily those of the Department of Health.

Submitted for publication January 11, 2012; revised February 13 and March 10, 2012; accepted March 10, 2012.

Disclosure: **D.P. Crabb**, None; **D.F. Garway-Heath**, Carl Zeiss Meditec (F, R), Heidelberg Engineering (F, R)

Corresponding author: David Crabb, Department of Optometry and Visual Science, City University London, London, EC1V 0HB, UK; d.crabb@city.ac.uk

Investigative Ophthalmology & Visual Science, May 2012, Vol. 53, No. 6

examination and, therefore, have had an impact on glaucoma management policy.[9] The notion that three tests per year should be performed to adequately measure visual progression is also supported by evidence from previous studies using computer simulation[10] and retrospective examination of a large VF data set.[11]

Some attention has been given to the prudent idea of varying the intervals between VF tests to optimize detection of progression.[12] In one novel approach, the length of the interval between subsequent tests depends on the outcome of previous test results.[13] Other alternatives to following a patient at evenly spaced intervals, such as every 6 months or annually, have also been suggested as a way of optimizing detection rates in clinical trials.[14] This report aims to determine whether estimates of progression, and rate of VF loss, would be improved if examinations were clustered at the beginning (baseline) and end of a predetermined observation period. The concept is motivated in part by well-established statistical theory on design of experiments for optimal slope estimation in linear regression.[15,16] In practice, the idea here is to set a follow-up period and then wait and see. This study, therefore, examines the hypothesis, using data from computer simulations, that wait and see offers advantages over evenly spaced test intervals in detecting progression. The report also aims to highlight the usefulness of this type of experiment (computer simulation) as a way of exploring how best to optimize the use of VFs when attempting to best detect progression in glaucoma and best identify, by way of improved estimates of rate of change, those patients that may or may not require intensified treatment.

Methods

In the computer simulations used in this study a VF was represented by a single number, equivalent to a VF mean deviation (MD). Change in MD is conventionally used as a marker for progression both in the clinic and in clinical trials; it is a summary measure of the overall reduction in VF sensitivity relative to a group of healthy age-matched observers. The follow-up period for all the simulation experiments was fixed to be 2 years. At baseline the VF was assumed to have an early to moderate glaucomatous defect with an MD value of −4 dB. For stable eyes, the true underlying MD was assumed to remain constant at −4 dB at every visit over the 2-year follow-up period. For progressing eyes, a 2 dB per year rate of loss was given to the MD, meaning that the MD rapidly deteriorates to −8 dB in a noise-free VF at the end of the follow-up period (Fig. 1A).

The next key step in the simulation was to add measurement variability to these true underlying values. Using the distribution of standard deviations (SDs) of MD from glaucomatous patients in a large longitudinal study,[17] Chauhan et al. (2008) defined patients exhibiting "moderate" measurement variability to have SD of MD of 1 dB.[8] For simplicity, this level of variability (SD = 1 dB) was adopted throughout all the simulation experiments, with a modification for progressing eyes. Variability was added to the MD recorded at each follow-up visit as a single number randomly sampled from a Normal distribution with mean 0 dB and SD of x dB. For stable eyes, x was fixed at 1 dB; for progressing eyes, x was modified to increase as MD worsens to mimic the real situation where VF measurement variability increases as the VF deteriorates.[18] There are few examples in the literature quantifying exactly how the SD of MD increases as the MD in an individual worsens. One study,[19] using frequency-of-seeing (FOS) data collected in patients, indicated that response variability, albeit at individual VF locations rather than MD, was reasonably well represented by the function, $\log_e(\mathrm{SD}) = \mathrm{A} \times \mathrm{sensitivity(dB)} + \mathrm{B}$, where the constants A and B are −0.081 and 3.27, respectively. So, since our simulation assumed that the initial VF has an MD of approximately −4 dB then the sensitivity values across the VF will be, on average, approximately 26 dB. A progression rate of 2 dB loss per year will give an MD of −8 dB after 2 years, meaning that the sensitivity values across the VF will be, on average in our example eye, approximately 22 dB. By substituting these two values (26 and 22 dB) into the equation, and then subtracting one result from the other, an estimate of the increase in SD of the sensitivity at a single deteriorating location is yielded (1.23 dB). The variability at a VF test location will naturally be higher than that in a summary measure such as MD So, following simple statistical sampling theory, this value of 1.23 dB was then divided by the square root of the number of locations in a typical VF ($n \approx 50$) to give an estimate of the expected increase in SD of the MD over the 2-year period at this approximate level of VF defect severity; this gave a value of 0.2 dB when rounded to one decimal place. Consequently, for the progressing eyes the model fixed the SD of MD at 0.9 dB at baseline and 1.1 dB at 2 years, with every 1 dB of loss in MD accompanied by a 0.05-dB addition to the SD of MD In this way the measurement variability in the simulation still assumes an average SD of MD of 1 dB, allowing useful comparison with the results from Chauhan et al. (2008),[8] but it also accommodates the characteristic of VF variability increasing over the follow-up period, as would be expected with real data from a progressing eye (Fig. 1B).

Four different follow-up schemes (timing of VF examinations) were investigated and were defined as: (1) 2 VF per year: evenly spaced. This follow-up scheme involves a "patient" being tested twice a year with evenly spaced, regular visits to the "clinic," requiring 5 VFs in total with one test at 0 (baseline), 6, 12, 18, and 24 months. (2) 2 VF: Wait and see. This follow-up scheme involves a "patient" taking two tests at baseline and two tests at 24 months, requiring only 4 VFs in total. (3) 3 VF per year: evenly spaced. This follow-up scheme involves a "patient" being tested three times per year with evenly spaced regular "visits" to the "clinic," requiring 7 VFs in total with one test at 0 (baseline), 4, 8, 12, 16, 20, and 24 months. (4) 3 VF: Wait and see. This follow-up scheme involves a "patient" taking three tests at baseline and three tests at 24 months, requiring 6 VFs in total.

For each of the four follow-up schemes, 10,000 different series of stable eyes (0 dB per year) and 10,000 different series of rapidly progressing eyes (2 dB per year) were simulated, with measurement variability added as described.

Linear regression of MD, giving estimated rates of loss (dB per year), was performed on each and every individual series. Eyes were classified as "progressing" if they satisfied standard criteria of a slope (rate of loss) of −1 dB or worse per year at the 5% level of statistical significance. Eyes not satisfying these criteria were considered stable. The proportion of "progressing" eyes fulfilling the criteria in the simulated progressing series gave an estimate of the power of the follow-up scheme (true-positive rate). In turn, eyes classified as "progressing" in the simulated stable series were defined as false-positive (FP) progression. With the evenly spaced follow-up interval schemes, linear regression was applied sequentially from the fourth test onward, so that progression could be detected at 18 or 24 months (or not at all) with the two tests per year scheme, or could be detected at 12, 16, 20, or 24 months (or not at all) with three tests per year scheme. Time to detection was therefore recorded for each "patient." This mimics the clinical scenario where VF series are tested for progression at each visit with a new VF. However, in both wait-and-see follow-up schemes progression could, of course, be detected only at 24 months, or not at all (Figs. 1C, 1D).

Estimating the rate of loss in patients is widely accepted to be an important part of the clinical management of glaucoma.[1,8] Therefore, the SD of all the 10,000 rates of loss estimates (slopes from linear regression in dB/year) was also calculated for each follow-up scheme. This acts as a measure of how well the rate of loss is estimated by the various follow-up schemes. A smaller SD of these slopes would suggest that the follow-up scheme is better for predicting true rate of loss.

All simulations and analyses were performed within the open-source statistical programming environment R (R Development Core Team, R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria, URL: http://www.R-project.org, 2008). All the R code used for this simulation is freely available from the authors.

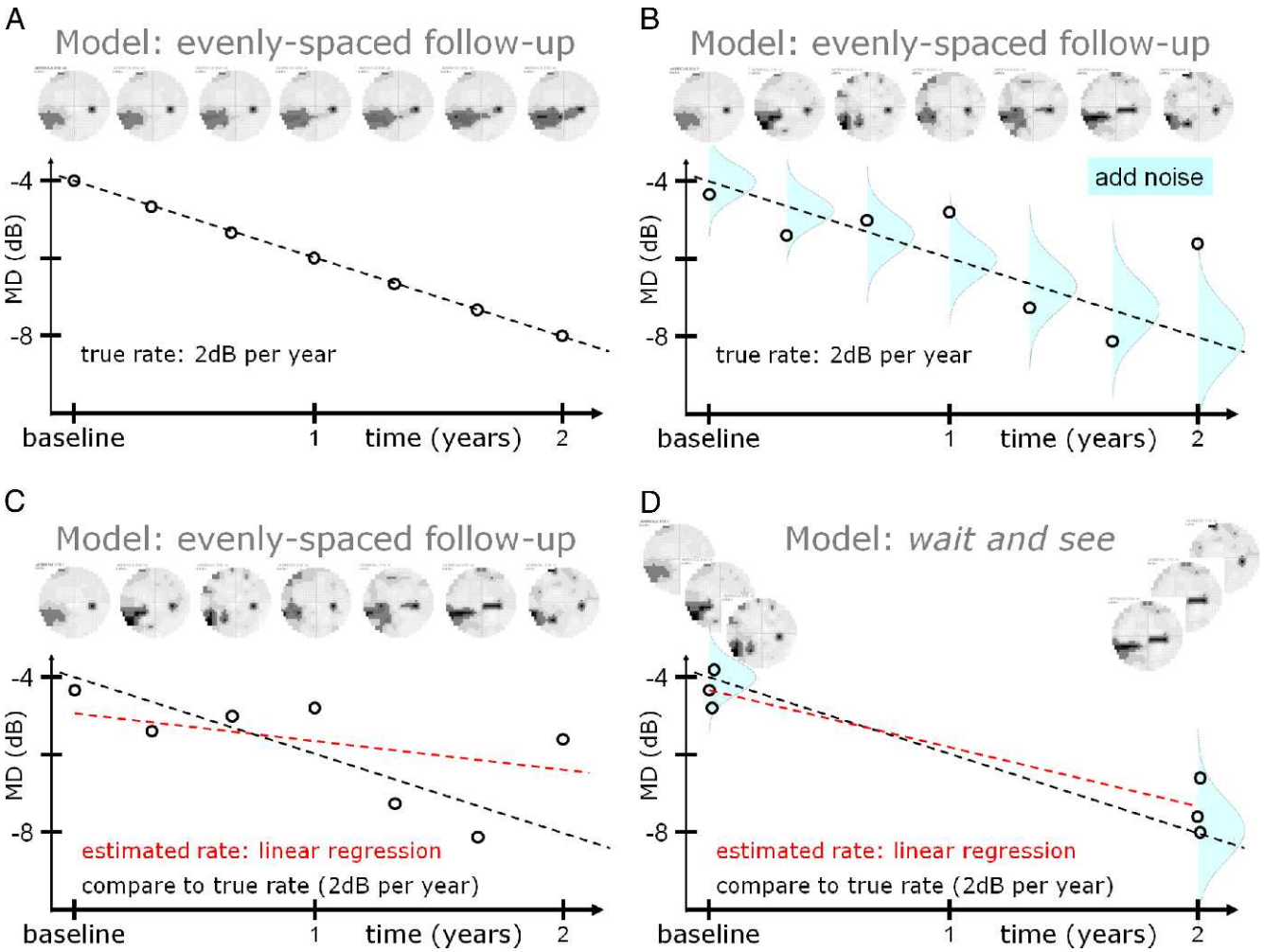

Figure 1. This schematic illustrates the steps in the computer simulation of a rapidly progressing eye. In (**A**) the eye is given rapid VF progression of −2 dB per year loss over a 2-year period (MD declines from −4 to −8 dB). The follow-up scheme is of regular, evenly spaced VF tests with 3 tests per year (7 VFs in total). The VF grayscale series shows the loss in the measurement variability-free eye. In (**B**) measurement variability is added to each observation, randomly sampled from a Normal distribution that centers on the true observation. The SD of this distribution increases as the VF sensitivity declines. The VF grayscale shows the progressing eye with measurement variability added, and is equivalent to the VF series that the clinician would observe; the VF is truly changing, but it is masked by measurement variability. In (**C**) the rate of loss is estimated using linear regression on the observed MD value over time and can be compared directly to the true rate of 2 db loss per year. If linear regression flags this as progressing (at least −1 dB per year at 5% statistical significance), then this series is declared a true positive result in the simulation. In (**D**) the equivalent process is carried out for the wait-and-see follow-up scheme. Note this requires one fewer VF test and, in this case the estimated rate of loss is closer to the true rate of loss than that observed with the evenly spaced follow-up (**C**).

Results

Summary statistics from the simulation experiments describe the "effectiveness" of each follow-up scheme. First, power of detection indicates how many rapidly progressing "patients" would be detected in the 2-year period and was calculated as the percentage of progressors detected or flagged from all 10,000 progressing patients. The overall false-positive (FP) rate for each follow-up scheme (% of 10,000 stable eyes falsely classified as rapid progressors) indicates the proportion of patients who would be wrongly diagnosed as rapid progressors when they are in fact truly stable. The average time to detection indicates how soon a rapidly progressing patient can be identified; this was calculated as a median time to event, including the censored values (those not detected in $n = 10{,}000$). These summary measures for the different follow-up schemes are shown in the Table.

Unsurprisingly, the power of detection was better with 3 VFs per year (evenly spaced) as compared with 2 VFs a year (evenly spaced), confirming the results previously reported in theoretical power calculations.[8] The wait-and-see follow-up schemes offered significantly better power of detection, coupled with better FP rates compared with the evenly spaced follow-up schemes (test of two binomial proportions; $P < 0.001$). With 3 VFs per year, the evenly spaced follow-up yielded a better average detection time (by 4 months, or one follow-up visit) when compared with the equivalent wait-and-see scheme (3 VFs). Yet the wait-and-see scheme correctly identified considerably more rapidly progressing patients (albeit at 24 months) and, by definition, requires one fewer VF test, when compared with the evenly spaced follow-up. FP rates were much better controlled with the wait-and-see schemes, because testing for progression is done once only (rather than after each VF test from the fourth onward) and the errors around the slope estimates are smaller. For example, as shown in the Table, with 3 VFs per year, the regular (evenly spaced) follow-up scheme gave an FP rate approximately 16 times worse than that provided by wait and see.

Table. Summary Statistics Describing the Effectiveness of Each Follow-up Scheme over the 2-Year Period

	2 VF Per Year: Evenly-spaced	2 VF: Wait and See	3 VF Per Year: Evenly-spaced	3 VF: Wait and See
Power to detect (%)*	58%	62%	82%	95%
False positives (%)†	2.9%	0.5%	5.9%	0.4%
Median detection time in months (IQR)	24 (18 to >24)	24 (24 to >24)	20 (16 to 24)	24 (24 to 24)
Number of VF tests	5	4	7	6
SD of rates of loss (dB per year)	0.63	0.51	0.56	0.41

IQR, Inter-quartile range. The 95% confidence intervals for the estimates of power to detect and false-positive rate (using the binomial distribution) based on $n = 10{,}000$ runs in the simulation are all smaller than (*) ±1% and (†) ±0.5%, respectively. Therefore, all the estimates of power to detect are significantly different and, in turn, the false-positive rates and SD of rates of loss for the wait-and-see schemes are significantly better than those for the respective evenly spaced schemes ($P < 0.001$).

Figure 2 shows histograms of the rate of loss (slopes from linear regression) for the $n = 10{,}000$ rapidly progressing patients generated for each of the four follow-up schemes. Comparison of the spread of these histograms gives a relative idea of how well the true rate of loss is estimated when using the different follow-up schemes. The narrower the distribution, the better the follow-up scheme is at estimating the true rate of loss. This can be quantified by the SD of these distributions (Table; also given in Fig. 2). The distributions were narrower and SD smaller for the wait-and-see follow-up schemes when compared with the evenly spaced follow-up schemes. True rates of loss were, therefore, better estimated with wait and see and the differences were highly significant (*F*-test of variances; $P < 0.001$).

Discussion

The statistics from this simulation experiment benefit from an illustrative discussion to aid clinical interpretation. Consider a glaucoma clinic that has 20 newly diagnosed patients who will progress rapidly in 2 years. Of course, at the point of diagnosis, without any monitoring, it is impossible to accurately identify such patients. If these patients are tested only twice a year, then the results from this report indicate that only approximately 12 of the 20 would be detected as having rapid VF progression (>2 dB per year loss) by 2 years of follow-up. If it is assumed that the clinic has sufficient perimetry resources for testing patients every 4 months (3 per year, evenly spaced) then this would result in approximately 16 of the patients being detected, a more acceptable figure for true positive detection in return for more frequent testing. If, however, the wait-and-see approach is adopted over the 2-year period, then nearly all of the rapidly progressing patients (19 of 20) would be detected at 2 years; an even better outcome than the evenly spaced follow-up and, furthermore, with one fewer test per patient. Now, it would be reasonable to assume that for every single case of a rapidly progressing patient there might be approximately 10 patients who are not rapidly progressing with VF loss that is relatively stable under treatment. This "illustrative" estimate is realistic, given findings from longitudinal and retrospective studies[3,20–23] (Heijl A, et al. *IOVS* 2008;49:E-Abstract 1155). For our imaginary clinic, these 200 patients are subjected to the same test frequency on follow-up as the rapidly progressing patients. The results of the simulation indicate that approximately 12 (5.9%) of these patients would be falsely detected as rapidly progressing if 3 VFs per year are used in an evenly spaced follow-up over the 2-year period. Because there are far fewer patients with true rapid progression on treatment as compared with relatively stable patients, the high frequency of testing (3 per year) would yield almost an equivalent number of correct (16 true positives) and incorrect (12 false positives) clinical decisions in the 2-year period. Arguments about the relative importance of true positive versus false positive progression identification are beyond this discussion, but these findings at least caution against the idea that simply increasing VF testing will be beneficial for all patients. The significant advantage of the wait-and-see approach, where analysis of results is done at one point in time, is false positive decisions are mainly averted. According to results of the simulation, around 1 in approximately 250 stable subjects would be falsely classified as being a patient with rapid progression when they are, in fact, stable

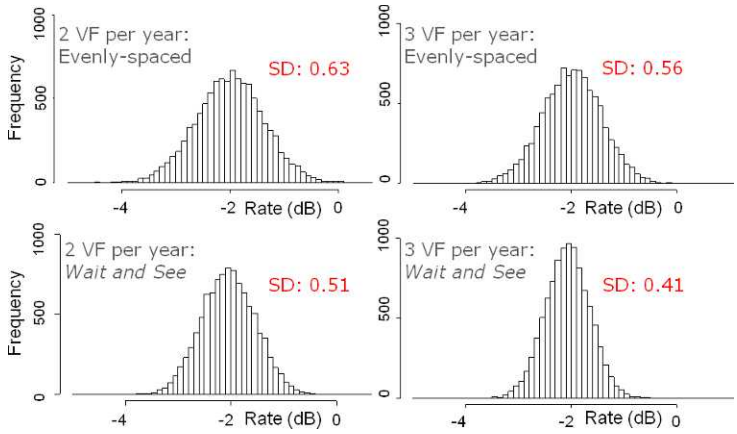

Figure 2. Histograms of the rates of loss (slopes from linear regression) for the $n = 10{,}000$ rapidly progressing patients generated for each of the four follow-up schemes. The wait-and-see scheme (6 VFs in total) has the narrowest distribution (and smallest SD) providing the best estimates of the true rate of loss. Note that even the wait-and-see scheme using just 4 VFs in total, estimates the rate of loss more efficiently (SD = 0.51) than the evenly spaced follow-up using 7 VFs in total (SD = 0.56).

under treatment. Adopting this approach would allow the clinician to have much more certainty that the flagged patient is actually rapidly changing, and can be more certain that this is not a false alarm, where the patient may undergo an unnecessary change in their treatment. Wait and see results in only a small delay in detection compared with evenly spaced follow-up; the results from this simulation indicate that the delay would be, on average, only approximately one visit (4 months), a modest difference in clinical detection time, especially given that one fewer VF test is required with wait and see.

The results from this report underline the futility of doing too few VF tests. For example, four or five VFs scheduled over 2 years will identify only approximately 6 of 10 rapidly progressing patients; annual testing would be worse. This result suggests that a real commitment should be made to do adequate testing in newly diagnosed patients; doing the odd VF test in every patient once or less per year is probably as bad as not doing it at all. Perhaps the key is to stratify patients into those that will benefit from more frequent testing or a certain testing follow-up scheme. Risk profiling patients to identify potentially fast progressors is difficult; risk factors for progression have been published,[24,25] but, as yet, there is no validated "risk calculator" for progression of manifest glaucoma and the accuracy of prediction from baseline factors is unknown. On the other hand, it is known that a major risk factor for going blind from glaucoma is advanced visual field loss at presentation.[26] Given limited VF testing resources, the clinician should attempt to identify patients at greatest risk of rapid progression and concentrate resources on these patients. All this needs to be done without falling below a safe frequency of testing that would identify unexpected progression in patients thought to be at low risk. Patients may also be stratified according to their ability to provide consistent VF measurements: for example, patients consistently returning noisy VFs may need a longer observation period before change or stability can be documented with confidence. This information about individual patient VF variability could be extracted from data from the "cluster" of tests that would be available at the beginning of the wait-and-see scheme. Of course, a prospective study in real patients would be needed to examine the advantages and problems of doing a cluster of tests over a few visits in terms of practicalities, patient acceptability, and learning effects. It would, for example, be interesting to examine the hypothesis that data might be more reliable if collected over a shorter period; testing in clusters at the beginning and end of a predetermined follow-up may obviate the need to continually reacquaint patients with SAP and may help to reduce the effect of long-term learning in VF testing.[27] Moreover, it might be reassuring for both patient and clinician if the wait-and-see scheme included an "alarm" VF at 1 year, with action taken if this measurement is dramatically worse than the baseline measures. This action would not assume rapid progression has occurred, but would result in acquiring confirmation VFs to maintain statistical power. Of course, this idea has not been tested here; the benefits of this adjustment, and others, would be best examined with further simulations or in a prospective study.

The simulation results draw attention to the problem of FP identification of VF progression, which has been shown to increase substantially with frequent periodic testing. This contradicts the perception that more frequent monitoring of the patient is in keeping with better clinical management. It is important not to make an FP diagnosis of progression in stable glaucoma because of the risks, side effects, and costs of unnecessary additional medical or surgical treatment. Reducing FP identification of disease progression in glaucoma management would also reduce the incidence of multiple medication changes: there is decent evidence that this may itself cause poor medication compliance and subsequent disease progression.[28,29] The wait-and-see approach offers more control over FP identification of rapid progression. This result is supported by evidence from a recent study on real data where test follow-up patterns were retrospectively modified in a similar fashion to the wait-and-see approach.[30]

There are limitations and observations to note about the results from this study. As with a previous report,[8] the results from this study relate only to the initial follow-up period for newly diagnosed patients. Moreover, these results are applicable only to VF testing, and not monitoring intraocular pressure (IOP). To date, IOP is the only evidential modifiable risk factor for disease progression. Therefore, IOP should still be monitored periodically, especially just after diagnosis and instigation of therapy, to ensure the patient has adequate IOP control and is adhering with therapy. It would, of course, be wrong to take an IOP measure (or measures) and send the patient away for 2 years. Another point, not tested in the computer simulations presented in this report but discussed elsewhere,[31] is to note the length of the observation period, rather than the periodicity of testing, which is critical to estimating the rate of loss with adequate precision, especially when making predictions of the course of disease in a particular patient.

Another observation about the results from this study is the assumption made that the rate of loss is linear and monotonic, rather than sudden or episodic. This is a reasonable assumption and broadly in line with current thinking about VF deterioration in glaucoma,[23] although there is no real evidence to dismiss the idea that for some patients the VF declines episodically or collapses in a short space of time. Nevertheless, it is noteworthy that in the computer simulation run in this study, if the 2 dB loss per year was assumed to be an "event" (a 4-dB episodic loss at some time point in 2 years) then the wait-and-see approach would always capture it. Although not formally tested here, this might not be the case with a regular follow-up when using linear regression to detect change because some observations at evenly spaced tests may have been made before the event occurred.

True status in any clinical series of data is uncertain, whereas in a modeled or computer-simulated series it is certain and thousands of tests can be carried out. Nonetheless, the accuracy and translation of the results are dependent on the assumptions and "input" to the model. For example, the model presented in this study considered only one level of measurement variability and one rate of loss (rapid progression, defined as 2 dB loss per year). This choice was deliberate to make the simulations simple to understand, accessible, and comparable to previous publications. The variability chosen is in line with expected "moderate" levels, but as shown previously, detection and FP rates, for example, worsen in those patients who demonstrate higher levels of measurement variability.[8] Yet looking at a range of patient variability, for example, would not affect the general conclusion about wait and see offering advantages over evenly spaced sequential follow-up. The wait-and-see approach is particularly applicable to clinical trials, where real gains in statistical power and improvements in estimating rates of loss in the different arms of the study could be made with no increase in resources. Such gains could be important in reducing sample sizes and numbers needed not to treat when a no treatment control arm is planned.

Progression behavior in glaucoma varies enormously, with some patients progressing rapidly to significant visual loss and the majority having slow or immeasurable deterioration. The heterogeneity of the disease behavior should be contrasted with the seemingly relative uniformity of disease monitoring, where almost all patients are reviewed at fixed intervals. So,

one aim of this report, via highlighting an alternative paradigm to testing at fixed intervals, is to stimulate others to investigate best practice for the diet of testing that should be used on newly diagnosed glaucomatous patients. This best use of resources should be high on the research and public health agenda because clinics strain under the numbers being monitored for this chronic condition, especially as the population age increases. For example, it has been estimated that glaucoma follow-up accounts for over one million visits to a clinic annually in the UK National Health Service alone.[32] Computer modeling experiments, as illustrated in this report, provide a very helpful first planning step before embarking on expensive trials to determine how best to monitor patients. There are other good examples of their use in modeling both VF[12,13,33–36] and structural measurements[37,38] in glaucoma over time.

In conclusion, this study demonstrates a follow-up scheme for monitoring VF progression as an alternative to evenly spaced examination intervals: wait and see will identify more patients with rapid VF progression with fewer FPs and one fewer VF. The concept of scheduling VF tests in this way is particularly applicable to clinical trials where VF progression is the functional marker for glaucomatous change. Computer simulation or modeling experiments, as reported here, are very useful for investigating and optimizing other follow-up schemes, before taking on the required costly and difficult prospective studies.

References

1. Caprioli J. The importance of rates in glaucoma. *Am J Ophthalmol*. 2008;145:191–192.
2. Collaborative Normal-Tension Glaucoma Study Group. Comparison of glaucomatous progression between untreated patients with normal-tension glaucoma and patients with therapeutically reduced intraocular pressures. *Am J Ophthalmol*. 1998;126:487–497.
3. Heijl A, Leske MC, Bengtsson B, et al. Reduction of intraocular pressure and glaucoma progression: results from the Early Manifest Glaucoma Trial. *Arch Ophthalmol*. 2002;120:1268–1279.
4. Chauhan BC, Mikelberg FS, Balaszi AG, et al. Canadian Glaucoma Study: 2. risk factors for the progression of open-angle glaucoma. *Arch Ophthalmol*. 2008;126:1030–1036.
5. Anderson DR, Drance SM, Schulzer M. Natural history of normal-tension glaucoma. *Ophthalmology*. 2001;108:247–253.
6. Heijl A, Bengtsson B, Hyman L, Leske MC. Natural history of open-angle glaucoma. *Ophthalmology*. 2009;116:2271–2276.
7. Friedman DS, Nordstrom B, Mozaffari E, Quigley HA. Glaucoma management among individuals enrolled in a single comprehensive insurance plan. *Ophthalmology*. 2005;112:1500–1504.
8. Chauhan BC, Garway-Heath DF, Goñi FJ, et al. Practical recommendations for measuring rates of visual field change in glaucoma. *Br J Ophthalmol*. 2008;92:569–573.
9. European Glaucoma Society. *Terminology and Guidelines for Glaucoma*. Savona, Italy: European Glaucoma Society; 2008 Available at: http://www.eugs.org/eng/EGS_guidelines.asp. Accessed July 22, 2011.
10. Gardiner SK, Crabb DP. Frequency of testing for detecting visual field progression. *Br J Ophthalmol*. 2002;86:560–564.
11. Viswanathan AC, Hitchings RA, Fitzke FW. How often do patients need visual field tests? *Graefes Arch Clin Exp Ophthalmol*. 1997;235:563–568.
12. Jansonius NM. Towards an optimal perimetric strategy for progression detection in glaucoma: from fixed-space to adaptive inter-test intervals. *Graefes Arch Clin Exp Ophthalmol*. 2006;244:390–393.
13. Jansonius NM. Progression detection in glaucoma can be made more efficient by using a variable interval between successive visual field tests. *Graefes Arch Clin Exp Ophthalmol*. 2007;245:1647–1651.
14. Schulzer M, Anderson DR, Drance SM. Sensitivity and specificity of a diagnostic test determined by repeated observations in the absence of an external standard. *J Clin Epidemiol*. 1991;44:1167–1179.
15. Daniel C, Heerema N. Design of experiments for most precise slope estimation or linear extrapolation. *J Am Stat Assoc*. 1950;45:546–556.
16. Gaylor DW, Sweeny HC. Design for optimal prediction in simple linear regression. *J Am Stat Assoc*. 1965;60:205–216.
17. Artes PH, Chauhan BC. Longitudinal changes in the visual field and optic disc in glaucoma. *Prog Retin Eye Res*. 2005;24:333–354.
18. Artes PH, Hutchison DM, Nicolela MT, LeBlanc RP, Chauhan BC. Threshold and variability properties of matrix frequency-doubling technology and standard automated perimetry in glaucoma. *Invest Ophthalmol Vis Sci*. 2005;46:2451–2457.
19. Henson DB, Chaudry S, Artes PH, Faragher EB, Ansons A. Response variability in the visual field: comparison of optic neuritis, glaucoma, ocular hypertension, and normal eyes. *Invest Ophthalmol Vis Sci*. 2000;41:417–421.
20. Smith SD, Katz J, Quigley HA. Analysis of progressive change in automated visual fields in glaucoma. *Invest Ophthalmol Vis Sci*. 1996;37:1419–1428.
21. The AGIS Investigators. The Advanced Glaucoma Intervention Study (AGIS): 7. The relationship between control of intraocular pressure and visual field deterioration. *Am J Ophthalmol*. 2000;130:429–440.
22. Broman AT, Quigley HA, West SK, et al. Estimating the rate of progressive visual field damage in those with open-angle glaucoma, from cross-sectional data. *Invest Ophthalmol Vis Sci*. 2008;49:66–76.
23. Bengtsson B, Patella VM, Heijl A. Prediction of glaucomatous visual field loss by extrapolation of linear trends. *Arch Ophthalmol*. 2009;127:1610–1615.
24. Leske MC, Heijl A, Hyman L, et al. Predictors of long-term progression in the early manifest glaucoma trial. *Ophthalmology*. 2007;114:1965–1972.
25. Friedman DS, Wilson MR, Liebmann JM, Fechtner RD, Weinreb RN. An evidence-based assessment of risk factors for the progression of ocular hypertension and glaucoma. *Am J Ophthalmol*. 2004;138(suppl 3):S19–S31.
26. Forsman E, Kivelä T, Vesti E. Lifetime visual disability in open-angle glaucoma and ocular hypertension. *J Glaucoma*. 2007;16:313–319.
27. Gardiner SK, Demirel S, Johnson CA. Is there evidence for continued learning over multiple years in perimetry? *Optom Vis Sci*. 2008;85:1043–1048.
28. Gurwitz JH, Glynn RJ, Monane M, et al. Treatment for glaucoma: adherence by the elderly. *Am J Public Health*. 1993;83:711–716.
29. Tsai JC, McClure CA, Ramos SE, Schlundt DG, Pichert JW. Compliance barriers in glaucoma: a systematic classification. *J Glaucoma*. 2003;12:393–398.
30. De Moraes CGV, Ritch R, Tello C, Liebmann JM. Modified visual field trend analysis. *J Glaucoma*. 2011;20:203–206.
31. Jansonius NM. On the accuracy of measuring rates of visual field change in glaucoma. *Br J Ophthalmol*. 2010;94:1404–1405.
32. National Institute for Health and Clinical Excellence. Glaucoma: diagnosis and management of chronic open angle

glaucoma and ocular hypertension. *Clinical guidelines*. 2009; CG85:4–37.

33. Gardiner SK, Crabb DP. Examination of different pointwise linear regression methods for determining visual field progression. *Invest Ophthalmol Vis Sci*. 2002;43:1400–1407.
34. Spry PG, Bates AB, Johnson CA, Chauhan BC. Simulation of longitudinal threshold visual field data. *Invest Ophthalmol Vis Sci*. 2000;41:2192–2200.
35. Vesti E, Johnson CA, Chauhan BC. Comparison of different methods for detecting glaucomatous visual field progression. *Invest Ophthalmol Vis Sci*. 2003;44:3873–3879.
36. Turpin A, McKendrick AM. What reduction in standard automated perimetry variability would improve the detection of visual field progression? *Invest Ophthalmol Vis Sci*. 2011; 52:3237–3245.
37. Vermeer KA, Vos FM, Lo B, et al. Modeling of scanning laser polarimetry images of the human retina for progression detection of glaucoma. *IEEE Trans Med Imaging*. 2006;25: 517–528.
38. O'Leary N, Crabb DP, Garway-Heath DF. An *in silico* model of scanning laser tomography image series: an alternative benchmark for the specificity of progression algorithms. *Invest Ophthalmol Vis Sci*. 2010;51:6472–6482.