Detecting Visual Field Progression in Glaucoma—
Using the Right Tools for the Job

Luke J Saunders¹, Richard A Russell² and David P Crabb³

¹. Full-time PhD student, Department of Optometry and Visual Science, City University London, UK, ². Post-doctoral Research Fellow in both Department of Optometry and Visual Science and NIHR Biomedical Research Centre for Ophthalmology, London, UK, ³. Professor of Statistics and Vision Research, NIHR Biomedical Research Centre for Ophthalmology, Moorfields Eye Hospital NHS Foundation Trust and UCL Institute of Ophthalmology, London, UK

Abstract
Monitoring disease progression is at the centre of managing a patient with glaucoma. This article focuses specifically on how visual field measurements from standard automated perimetry (SAP) can be monitored over time. Various options for analysis on the Humphrey and Octopus perimeters are discussed, from summary indices to event and trend-based analyses, and their respective merits and flaws are evaluated. It is strongly recommended that quantitative analysis methods and software are used in assessing progression, as variability in threshold measurements makes detecting true deterioration non-trivial. Recommendations on the frequency of visual fields that should be taken per year are also discussed. The article concludes by putting the spotlight on new research being undertaken to improve the methods of measuring and predicting progression, as well as relating visual fields to patient visual disability and quality of life.

Keywords
Glaucoma, visual field (VF), progression analysis, standard automated perimetry

Gliacoma is a set of eye conditions in which the optic nerve head (ONH) and retinal ganglion cells are damaged, resulting in the visual field (VF) of the sufferer being reduced. Part of what makes glaucoma so dangerous is that the rate of VF impairment is typically slow and that it usually begins by affecting the peripheral vision of a sufferer. Consequently, many patients do not report the condition until very late when the amount of VF lost can seriously undermine their quality of life (e.g. by having their driving licence revoked or being at increased risk for a fall). Naturally, most methods of glaucoma treatment involve reducing the intraocular pressure and this will require adapting based on the rate of disease progression, which makes monitoring VF progression important. Information on VF progression is also discussed. The article concludes by putting the spotlight on new research being undertaken to improve the methods of measuring and predicting progression, as well as relating visual fields to patient visual disability and quality of life.

Measuring the Functional Progression of Glaucoma
One of the most important methods of diagnosing glaucoma and monitoring its progression is through measuring the functional progression of the disease. Structural methods, looking at the health of the ONH, can also be used but these are not the focus of this article. Perimetry is the means by which the VF of a patient is mapped and the only means of measuring functional progression. The Humphrey® Field Analyzer (HFA; Carl Zeiss Meditec, Dublin, CA) and the Octopus (Haag-Streit, Köniz, Switzerland) are commonly used perimeters, especially in a tertiary or referral setting where the goal is to monitor VFs in patients with glaucoma or at risk for developing glaucoma, so this article will focus on using these instruments. Standard automated perimetry (SAP) has been the gold standard to measure VF since its introduction in the 1980s, but there are a number of other automated perimeter methods that have been developed to try and displace it, including short wavelength automated perimeter (SWAP), pulsar (or flicker perimetry), and frequency doubling perimetry (FDP or FDT). However, research is ongoing and, for now, SAP remains the primary method for detecting VF progression and is hence the subject of this article. As with structural methods, alternative functional methods are helpful to be used alongside rather than instead of SAP.

The full-threshold method to measure thresholds in SAP employs a ‘staircase’ technique. This method provides measurement of defect severity at each location in the VF but can take quite a long time (about
Detecting Visual Field Progression in Glaucoma—Using the Right Tools for the Job

15 minutes per eye), during which it can be difficult to keep a patient’s attention. Thus, faster methods have been devised in order to reduce test time. One such technique for the HFA is the Swedish interactive threshold algorithm (SITA) Standard method. This strategy seeks to reduce test time by incorporating the knowledge that the threshold at each point is unlikely to be independent of neighboring locations. The large advantage of SITA Standard is that it can halve the duration of testing for many patients. An equivalent, though different, strategy for the Octopus perimeter is known as the ‘Dynamic’ strategy, which shares many of SITA Standard’s advantages. Other even-faster methods have been devised by researchers in order to enhance the speed of testing further (such as SITA Fast and Tendency-Orientated Perimetry for the Humphrey and Octopus, respectively), but these may sacrifice sensitivity in identifying VF defects and there is some evidence to show the results are less repeatable.

A point for consideration when reviewing perimetry results is the means of assessing how reliable these are, as poor results lead to misdiagnosis. All of the tests described above have methods of evaluating the reliability of the test itself through looking for false negatives, false positives, and fixation losses. These measures give an indication of patient attentiveness, how ‘trigger-happy’ the patient is and the patient’s fixation performance, respectively. Major clinical trials have used these measures to assess VF reliability, yet the criteria applied to them remain arbitrary and vary between trials. For instance, for the AGIS and CIGTS trials, a scoring system was used that meant that patients could theoretically pass with false positive or negative rates in excess of 33 %, while the EMGT did not look at the false negative rate at all. In addition, other reliability criteria such as the total number of questions asked (longer tests imply greater uncertainty in measuring the threshold) and short-term fluctuation values have been used. There is no single ‘perfect’ reliability index in assessing progression but if a VF is unreliable then this will hinder the ability to detect progression. Furthermore, there is evidence to suggest that the reliability indices do not provide accurate insight about a patient’s performance. Instructions to the patient, correction of spherical ammetropia, and patient attention all have a significant bearing on the result and reliability indices should not be relied upon exclusively. Furthermore, the psychophysical nature of the tests means that learning effects need to be taken into consideration, as patients often improve in their ability to undertake perimetric testing with experience. Although seemingly wasteful, discounting initial VFs remains good practice when assessing glaucoma progression. There are several good publications about the practicalities of good VF testing with SAP and interpretation of results, a comprehensive one being by Henson. However, even after following all the recommendations for good practice, the clinician is often left with variable VF measurements upon which to make a decision. It is this ‘noise’ in the measurements of VF thresholds, which makes measuring progression far from straightforward, particularly in areas of damage, where variability is even greater.

Summary Methods of Detecting Functional Progression in Glaucoma
The output produced from SAP can be confusing as it contains a huge amount of data and information, and it is not always particularly obvious how large changes are from one VF assessment to the next. In spite of this, many clinicians judge progression ‘manually’ using their experience in comparing SAP printouts. However, there is good evidence that this...
Furthermore, it has already been shown that, when clinicians use software, their concordance in judging progression is mediocre at best.\textsuperscript{15,16} Furthermore, it has already been shown that agreement between clinicians for judging progression is mediocre at best.\textsuperscript{15,16} Thus, it is vital that quantitative methods are utilized for measuring progression. Detecting VF progression requires the thorough understanding and utilization of the available software and analysis methods to facilitate this task and these will now be reviewed.

Global indices are often utilized to assess the VF. An important feature of any such statistic is that it can be put into some context; consequently, every perimetry machine has a normative database of VF thresholds to compare measured VF thresholds against at a given location and for a given age, so that a significant loss can be determined. At each point, the total deviation (TD) value represents the difference between the measured threshold and the expected VF sensitivity for a healthy eye of the given age. Summary measures such as the mean defect or mean deviation (MD), take an average of all TD values in the test to produce a single statistic of global VF sensitivity compared with that of an aged matched person with ‘normal’ vision.

The problem with using MD as a measure of VF damage is that it is only useful in terms of summarizing the defectiveness of the whole VF, when, in glaucoma, the size and position of localized damage are likely to be of most interest (see Figure 1). The TD plot partially resolves this concern, but it does not differentiate the effects of glaucoma from overall sensitivity loss associated with worsening media opacity from cataract. The pattern deviation (PD) plot was devised to assess the location of defects, while also separating what is likely to be glaucomatous from general VF worsening caused by cataract. Pattern standard deviation (PSD) then measures the amount of variability in the PD values and so attempts to resolve focal from diffuse loss. PSD is not a good indicator of overall damage in advanced disease, though it can still be helpful when used alongside the MD.\textsuperscript{17} However, neither PSD nor MD provide spatial information about the VF defect.

The Visual Field Index (VFI) is a relatively new summary measure that seeks to quantify glaucomatous damage in newer models of the HFA. VFs with no discernable defect are scored 100 % with 0 % signifying perimetric blindness. For each defective point in the pattern or TD plot, the VFI computes a weighted average, operating on the principle that the very central part of the VF is of highest importance and so is weighted more. The VFI attempts to measure only damage from the PD values rendering it some immunity to the confounding effects of cataract and it prioritises the very central VF more, therefore, giving a better estimate of actual functional loss.\textsuperscript{18} Interestingly, however, there is evidence to point out that the VFI overestimates what would be attributed by expert opinion on the percentage of functionally useful VF remaining of an eye; in

\begin{figure}
\centering
\includegraphics[width=\textwidth]{figure2}
\caption{A Demonstration of the Differences Between (A) Event and (B) Trend-based Analyses for One Point in Consecutive Visual Fields}
\end{figure}
short, the VFI may be somewhat misleading in representing how well a patient can see.19 Thus, though the VFI is useful, MD should still be considered when evaluating VF loss. A better index, in terms of VF disability, could potentially be gleaned by combining measures from two eyes,20,21 but no specific measures combining monocular VFs have yet been tested and clinically accepted. The integrated VF (also known as the best sensitivity method), which takes the best sensitivities from either eye for each point, is one such method that has so far had promising results compared with conventional binocular tests unavailable to the clinician.22-24

### Method

<table>
<thead>
<tr>
<th>Method</th>
<th>Method of Attaining Baseline</th>
<th>Method of Defining Progression</th>
<th>Advantages</th>
<th>Disadvantages</th>
<th>Correction for Cataract?</th>
<th>Method Type</th>
<th>Rate of Progression Calculable?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear regression of MD values</td>
<td>No consensus, but at the very least three VFs are required</td>
<td>No consensus, but EyeSuite for the Octopus perimeter defines progression if the previous six MDs are significantly progressing (p&lt;0.5 %)</td>
<td>• Simple</td>
<td>• No spatial consideration of the data</td>
<td>No</td>
<td>Global Summary Trend Analysis</td>
<td>Yes</td>
</tr>
<tr>
<td>Linear regression of VFI values</td>
<td>Five VFs over 3 years required for initial trend in Humphrey Field Analyser software</td>
<td>Assesses significance of slope, and relates rate of progression to how much VFI patients will lose in the next 5 years</td>
<td>• Simple</td>
<td>• No spatial consideration of data</td>
<td>Yes</td>
<td>Global Summary Trend Analysis</td>
<td>Yes</td>
</tr>
<tr>
<td>AGIS Method</td>
<td>One VF</td>
<td>A decline in score from baseline equal to four ‘AGIS units’ in three consecutive tests</td>
<td>• High specificity;0,12,28 Score testing based on real patient data</td>
<td>• Poor sensitivity;0,12,28 Cannot determine spatial characteristics of progression</td>
<td>No</td>
<td>Pointwise/scoring event analysis</td>
<td>No</td>
</tr>
<tr>
<td>CIGTS Method</td>
<td>Two VFs</td>
<td>A decline in score from baseline equal to three ‘CIGTS units’ in three consecutive tests</td>
<td>• Fast;0,12,28 High specificity;0,12,28 Score testing based on real patient data</td>
<td>• Low sensitivity;0,12,28 Cannot determine spatial characteristics of progression</td>
<td>No</td>
<td>Pointwise/scoring event analysis</td>
<td>No</td>
</tr>
<tr>
<td>GCP Method</td>
<td>Two VFs</td>
<td>‘Likely progression’ defined as a reduction in sensitivity (below normal limits) from baseline for ≥3 separate VF points in three consecutive tests</td>
<td>• Fast;0,12,28 High sensitivity and specificity;12,28 Normal limits defined using real (stable) patient data</td>
<td>• Cannot detect progression rate</td>
<td>Yes</td>
<td>Pointwise trend analysis</td>
<td>No</td>
</tr>
<tr>
<td>PLRA</td>
<td>No consensus, but at least three VFs are required</td>
<td>No consensus, but usually a statistically significant (p&lt;5 %) decrease of 1 dB per year for at least three separate VF points</td>
<td>• High sensitivity;0,12,28 High specificity;0,12,28</td>
<td>• Long time period required</td>
<td>No</td>
<td>Pointwise trend analysis</td>
<td>Yes</td>
</tr>
</tbody>
</table>

AGIS = Advanced Glaucoma Investigative Study; CIGTS = Collaborative Investigative Glaucoma Treatment Study; GCP = glaucoma change probability; MD = mean deviation; PLRA = pointwise linear regression analysis; VFI = visual field index.
Glaucoma

Figure 3: A Simplified Graphical Representation of Table 2 from the Paper by Chauhan et al.4

The graph shows the number of visual fields (VFs) per year required to have an 80 % probability of successfully detecting a raw change in mean deviation (MD) in a given number of years for a patient with average variability. For example, to have an 80 % chance of detecting a change of 1 dB over 5 years, four VFs per year are required. By contrast, to catch a fast progressor who has suffered a 4 dB loss over 3 years, three VFs per year are required to have a 4 in five chance of detecting it. The greater the size of the change and the number of VFs that are taken, the higher the likelihood that the change is detected.

Overall, though summary measures are useful in terms of having a singular measure representing how badly an individual’s VF has been degraded by glaucoma, a universal problem for all of these measures is that they waste data and ignore important spatial information. Furthermore, it is difficult to tell how changes in these measures relate to visual disability. For example, what exactly does a drop in VFI from 100 to 97 % mean? Criteria must be devised in order to indicate whether glaucomatous progression is taking place at a dangerous rate or not. Pointwise scoring criteria (examining each VF location separately) that amount to counting defective points have been used for determining progression, though usually only in clinical trials, as they seem to have acceptable diagnostic specificity (probability of diagnosing stable eyes as nonprogressing). Precise details on the methods used for evaluation of VF progression in some major clinical trials can be found in the papers referenced here. However, better methods for detecting change, namely event and trend-based analyses, have now been incorporated into software for the analysis of progression.

Event and Trend-based Analyses

There is often much debate between whether to use event-based analyses or trend-based methods. The main difference between the two is to do with how they treat measurements taken between the first and last VF, which gives them their specific properties. Event-based analysis involves taking a baseline and comparing every subsequent test against this reading (see Figure 2a). Any significant difference between the baseline and latest reading is considered to be due to progression.

One method of this kind is the ‘glaucoma change probability’ method (GCP), which is sometimes called ‘glaucoma probability analysis’ and is available from Statpac 2 software for the HFA. In short, the pointwise VF sensitivities from an average of two baseline readings are compared with subsequent VF tests using the GCP map (GCPM), which considers the amount of ‘normal’ variability one would expect for the baseline values. Points that show significant deviation from baseline in three consecutive tests are marked as progressing (see Figure 2a). Heijl et al. used an adapted version of the GCPM for the EMGT. The GCP method is a sensitive method for detecting glaucomatous progression and there is evidence to suggest that results from this analysis have good agreement with expert opinion on progression.9

Trend-based analyses (see Figure 2b) are an evolving process in which all VFs are analyzed using linear regression to assess the rate and significance at which the measurement is changing over time. Trend-based analyses can be applied to summary measures such as VFI or MD, and are performed in the Statpac (HFA) and EyeSuite (Octopus) Progression Analysis software. Recently, this approach has been advocated, using the VFI in particular,9 but linear regression of summary measures has been criticized as being relatively insensitive to detecting VF progression.20

An alternative method is Pointwise Linear Regression (PLR), implemented on PROGRESSOR (PROGRESSOR Medisoft, Leeds, UK), EyeSuite (Haag-Streit, Kôniz, Switzerland) and Peridata (Peridata Software GmbH, Huerth, Germany) software, which detects the rate of progression of each point in the VF. A common statistical criterion used to assess whether progression is significant for a single point is a rate less than −1 dB/year (p<0.05). Other criteria are examined in detail by Gardiner and Crabb.22

As a result of using all the previous fields in its diagnosis, trend analysis has statistical advantages and disadvantages compared with event-based analyses. First of all, event-based analyses generally require fewer VFs and less time to produce definitive results, so may detect rapid deterioration in the VF more quickly than trend analysis.27 However, trend analysis is an important diagnostic of progression because it estimates the rate of VF progression whereas event analyses focus only on a significant magnitude of difference between two values. Furthermore, trend analyses have been shown to have higher diagnostic sensitivity than event analyses.27 On the other hand, trend analyses often take longer to establish progression and the inference changes with each new measurement. For example, in Figure 2b, progression would be diagnosed after 3.5 years, but after 4 years the eye is again considered as not progressing (though after 4.5 years progression is once again diagnosed). For all their perceived differences, event- and trend-based methods are quite similar; the resolution of the analysis is at a single point and they can, with the aid of graphics used in software, highlight the spatial location of change that an experienced clinician can act upon. Still, there remains no general consensus on the number of points that constitutes ‘real’ progression, whether there should be a requirement for contiguous points to show this behavior or whether they should be maintained in subsequent fields. For a comprehensive review summarizing the evidence base on the different methods for assessing glaucomatous VF progression readers are directed to a review by Ernst et al.,28 but a practical summary of the differences between all the methods discussed is given in Table 1.

The Frequency of Visual Field Testing

Another important topic regarding how progression is detected refers to how regularly VFs should be taken. Chauhan and colleagues make suggestions based on the power (the probability of correctly diagnosing a progressing patient) associated with the number of tests taken. If too few VFs are taken of a given patient, detecting glaucomatous progression is unlikely due to the inherent variability of VF measurements. A more accurate estimate of the rate of progression is obtained by taking more readings because the underlying signal is more likely to be detected in the variability. For instance, it would be foolish to conclude that there are no fish in a lake based on one unsuccessful fishing trip. The
Detecting Visual Field Progression in Glaucoma—Using the Right Tools for the Job

Figure 4: Screenshots from a) Statpac 2’s Glaucoma Probability Analysis for the Humphrey® Field Analyzer, b) EyeSuite Analysis Software for the Octopus Perimeter, c) PROGRESSOR, and d) Peridata’s Boxplot Trend Analysis

probability of there being no fish depends on how long the fishing trip lasted and how many fish were in the lake. The duration of the fishing trip is analogous to sample size—if you collect more data you have a greater power to detect change. The number of fish in the lake is analogous to the effect size—you have greater power to find a large effect than a small one. Chauhan et al. suggest, as a minimum, that six VFs should be taken in the patient’s first 2 years of monitoring, before choosing the number of subsequent tests per year on the basis of progression rate and time-scale thereafter (see Figure 3). Of course, taking more VFs in a period of time will increase the number of false positive diagnoses, i.e. decrease specificity; though specificity can be improved by using more conservative progression criteria.

Gardiner and Crabb found, using simulation, in highly variable eyes with thresholds decreasing by 2 dB per year, undergoing three VF tests per year was the optimum number in terms of sensitivity and specificity. However, Nouri-Mahdavi et al. point out that, in clinical practice, it is often difficult, time-consuming, and costly to carry out so many VFs each year and hence recommend measuring VFs once every 6 months. However, there is a substantial loss in sensitivity by taking less than three VFs a year. Some attention has been given to the prudent idea of varying the intervals between VF tests to optimize detection of progression. One novel approach that should appeal to the interested reader is to adopt an approach that varies the length of the interval between subsequent tests depending on the outcome of previous test results. In conclusion and in order to shift the discussion away from a statistical one, it should be emphasized that a real commitment should be made to carry out an adequate amount of testing in newly diagnosed patients; performing the odd VF test in every patient, say once or less per year, is probably as bad as doing no VF testing at all. The key is, perhaps, to stratify patients into those that will benefit from more frequent testing. However, accurate risk profiling of progression awaits further research.

Software for Glaucoma Monitoring

The importance of using the correct software to aid diagnosis of VF progression is stressed again here. It should be noted that software should not be used in isolation from either clinical judgment or other measures and risk factors, but it is essential in aiding the clinician to make the appropriate assessment of VF progression. Statpac 2 for the Humphrey perimeter contains the means of performing GCP analysis as well as all the summary indices described earlier in this review (though no PLR analysis). The EyeSuite Progression Analysis software for the Octopus, meanwhile, has its own range of output including all the summary measures described above bar the VFI, as well as a program combining structural and functional defects, trend analyses of MDs, and PLR, but does not yet feature a GCP analysis program. PROGRESSOR implements PLR analysis using color-coded bar charts for each VF threshold, while Peridata allows PLR and also has the ability to display colored diagrams of defect depth and 3D hills of vision (both can be used alongside the HFA and Octopus perimeters) (see Figure 4).

Discussion

This review has summarized the main approaches for assessing glaucomatous progression and the following conclusions can be made. First, SAP remains the current gold standard for measuring functional...
Many researchers have utilized 2003;44:2613–20. event-based methods are useful for monitoring and detecting progression, and should be employed according to the clinical objective. Third, reliability indicators in addition to those provided by perimeters need to be taken into account when assessing whether an obtained VFI is sufficiently accurate. Finally, an adequate number of VF examinations should be taken in the first 2 years of monitoring a patient in order to maximize the probability of detecting glaucomatous progression.

Hopefully, in the near future, new methods could yet allow for more accurate monitoring of glaucoma progression. For example, novel research aims to determine how structural defects can be better linked to functional ones to aid diagnoses with promising results thus far. Incorporating Bayesian statistics into monitoring VF progression—using prior knowledge of patients with glaucoma or structural information—is another interesting research topic.43-46 Many researchers have utilized computer simulations in order to compare methods.27,32-35 This provides a method for benchmarking different techniques in the absence of a true gold standard for VF progression, so the future may see this method used increasingly. New modeling methods that navigate the problem of detecting VF progression better than standard linear modeling are also being investigated.45,47

Finally, little is known about how VF defects, at different stages of glaucoma, affect patients’ abilities to perform everyday visual tasks. Being able to link the measurements taken in the clinic to what patients visually ‘can’ and ‘cannot do’ would be enormously helpful. In particular, the VF component for legal fitness to drive is an area that is drawing much investigation28,48-50 due to the fact that this is closely related to an individual’s quality of life.50 In this respect, the future benchmark for the success of new glaucoma treatments to negate progression could be aligned to a measurable reduction in visual disability rather than imperceptible changes on a clinical chart. One conclusion that we can be certain of is that, for the moment, detecting VF progression is nontrivial and quantitative tools are essential to aid clinical decisions.11

This article was originally published for the European audience in: European Ophthalmic Review, 2013;7(1):20-6