

Mean deviation fluctuation in eyes with stable Humphrey 24-2 visual fields

CL Tattersall, SA Vernon and GJ Menon

Abstract

Aim To describe the expected fluctuation in the mean deviation (MD) scores on a large, long-term series of stable visual field reports for particular grades of defect in order to give clinicians an aid to the correct diagnosis of glaucomatous progression.

Method Visual field reports of subjects with five reliable consecutive Humphrey 24-2 visual fields, recorded over a period of at least 3 years, were scored using the Advanced Glaucoma Intervention Study (AGIS) system. The AGIS scores of the first and last visual fields were required to be identical.

Results A total of 202 eyes from 202 patients were used in the study, with a total of 1010 visual fields being used in the analysis. Visual fields with no defect (AGIS score 0) had a 99% confidence interval (CI) of 0.3 dB, mild defects 0.4 dB, moderate defects 0.8 dB, severe 1 dB, and 1.3 dB for end-stage defects when considering variation of MD scores. Using a one-way ANOVA incorporating all stages showed very little fluctuation throughout the series ($P = 0.96$). The correlation between the CI and grade of field defect showed a good positive correlation ($r = 0.7$, $P = 0.0003$) indicating an increase in CI as field defects worsen.

Conclusion When considering a series of reliable visual fields in a stable eye, one should expect only very little fluctuation in the MD, indicating that an increase in the MD beyond that of the 99% CI described may suggest progression. The reliability indices of the visual field test should be regarded as a primary consideration when assessing visual fields.

Eye (2007) 21, 362–366. doi:10.1038/sj.eye.6702206; published online 20 January 2006

Keywords: mean deviation; visual fields; fluctuation; glaucoma

Introduction

Computerised visual field testing remains after many years the benchmark for monitoring glaucoma, with the Humphrey analyser established as the gold standard in the UK. It is however, a subjective test and is dependent upon such factors as the patients' reliability or fatigue.¹ Visual fields can therefore vary from test to test, despite any actual defects remaining stable. This adds to the complexity of glaucoma management.

Studies in which this aspect of glaucoma management has been examined^{2–5} have indicated that a certain amount of fluctuation in visual fields may exist. In these studies the identification of intertest fluctuation used complicated analysis of various indices on the visual field tests, related to field tests separated by between 1 week and 5 years, and use no more than three visual fields in a series. Boeglin *et al*⁶ published a range for the mean deviation (MD) in stable visual fields, however, this was not separated for stage of glaucoma, thus giving a broad range of fluctuation.

The identification of stable visual fields in a series is problematic, and has been subject to much study, with the development of software packages such as PROGRESSOR and STATPAC 2.⁷ These software packages are designed to identify progression, which they have been shown to do with good accuracy.⁸ However, for many patients', differing visual field strategies have been employed throughout a series of visual fields, or visual field software packages are not in use. In order to manually identify stability, the Advanced Glaucoma Intervention Study (AGIS) system⁹ was employed in this study, which is a quantitative assessment of visual fields and has become a benchmark for the manual assessment of visual fields. The AGIS system combined with a strict reliability

Directorate of
Ophthalmology, Queen's
Medical Centre,
Nottingham, UK

Correspondence:
CL Tattersall,
Department of
Ophthalmology,
Royal Glamorgan Hospital,
NHS TRUST,
Ynysmaerdy,
Llantrisant CF72 8XR, UK
Tel: +44 1443 443 443;
Fax: +44 1443 443 443.
E-mail: clt06@aol.com

Received: 21 June 2005
Accepted in revised form:
18 October 2005
Published online: 20 January
2006

criteria¹⁰ will as accurately as possible determine stability within a series of visual fields which incorporates various field strategies not used with, or incompatible with software assessments.

In a time-sensitive environment, such as in many ophthalmic clinics, an accepted and quick assessment of glaucomatous defects for clinicians can be made using the MD, which is available on any test strategy. The MD corresponds to the mean elevation or depression of the patients' visual field compared to a normal reference field.¹¹ The calculation of the MD takes into account age-corrected normal values and the normal degree of variance at each of the 54 test locations used in a Humphrey 24-2 visual field.¹²

The aim of this study is to describe the expected fluctuation in the MD scores in a large, long-term series of stable visual fields for particular grades of defect in order to give clinicians an aid to the correct diagnosis of glaucomatous progression, which can be used in an everyday clinic setting.

Methods

The visual fields of patients with a series of five Humphrey 24-2 visual fields over a period of at least 3 years were analysed. Patients were selected by reviewing clinic notes collected consecutively over a period of 12 months. Glaucomatous defects were based on the AGIS visual field score. In identifying a suitable series of visual fields the AGIS scores (ranging from 0 to 20) of the first and last fields needed to be identical. However, to subdivide the analysis the broader levels within the AGIS system were used (nil, mild, moderate, severe, end stage). These levels incorporate the numerical scores as shown in Table 1. All five visual fields in a series were required to be reliable, where reliable was defined as both the false-positive and false-negative indices being less than 20%.¹⁰

A phenomenon known as the 'learning effect' has been identified^{13,14} whereby the patients' baseline visual field is worse than subsequent tests. The learning effect was a consideration in the study as using these initial tests may have resulted in a false-positive defect being compared to actual glaucomatous progression or the development of

cataract, and not stable fields. Therefore, if the patients' tests were to be included in the study, the second field test could not show an improvement ≥ 2 dB in the MD over the first.

Only one eye per patient was used in the study, and were all scored by one of the research team (CT). Of the visual fields, 20% were chosen at random and independently scored by a senior trainee ophthalmologist (GJM) again using the AGIS system in order to ensure a valid identification of visual field scores.

Results

An estimated 700 sets of medical notes were considered for the study, with 300 of these having a series of visual fields and therefore being closely reviewed for inclusion. In all, 74 sets of fields from the original 300 were excluded on the basis that the first and last were not at an identical stage. A total of 16 were excluded due to the reliability of the series being in question, with eight excluded due to a learning effect being evident. In all, 202 eyes were therefore used in the study, with a total of 1010 visual fields being available for analysis. The number of eyes and visual fields available for analysis in each AGIS level is shown in Table 2. The mean time span of a series of five fields was 5.8 years (SD 2), with the mean age of patients being 71 years (SD 11.3) at the time of the final visual field. Of the patients, 109 were male.

To ensure no bias in the results due to the age of the patients, or time since the first visual field, the groups (based on level of field defect) were assessed using an ANOVA test. This showed no significant difference between the groups with regard to age ($P = 0.17$) or time ($P = 0.45$). All AGIS scores for visual fields scored independently by CLT and GJM were in agreement (Kappa score: 1).

The average MD and 99% confidence interval (CI) for each level is shown in Figure 1. Visual fields with no defect had a CI of 0.2 dB, mild defects 0.3 dB, moderate defects 0.6 dB, severe 0.8 dB, and 1 dB for end-stage defects. To confirm the amount of fluctuation found, the

Table 1 Levels in AGIS system

Score	Level of glaucoma
0	None
1-5	Mild
6-11	Moderate
12-17	Severe
18-20	End stage

Table 2 Fields/eyes for analysis

Field defect	Number of eyes	Number of fields
None	68	340
Mild	71	355
Moderate	34	170
Severe	17	85
End stage	12	60
Total	202	1010

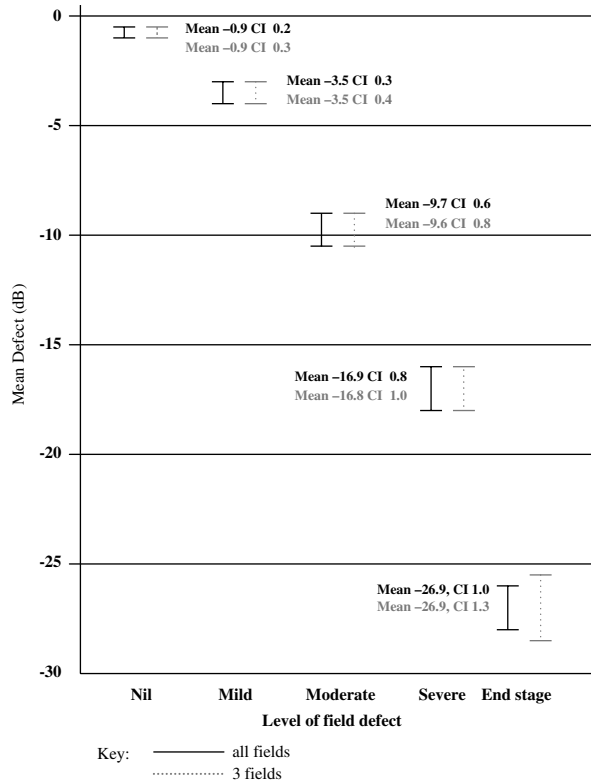


Figure 1 99% CI and mean.

series of five fields were compared using an ANOVA test incorporating all levels, and showed very little fluctuation throughout the series ($P = 0.99$). Although little fluctuation across a series was indicated, the correlation between the CI and level of field defect was analysed and showed a good positive correlation ($r = 0.7$, $P = 0.0003$) indicating an increase in CI as the field defect worsens.

As the first and last visual field in each series were known to be similar and reliable, these were then excluded, leaving the three remaining fields in each series. This analysis therefore gave an assessment of fluctuation within a stable series of visual fields. The 99% CI for each level is also shown in Figure 1. Visual fields with 'no defect' had a CI of 0.3 dB, 'mild defects' 0.4 dB, 'moderate defects' 0.8 dB, 'severe' 1 dB, and 1.3 dB for 'end-stage' defects. Again using an ANOVA test incorporating all levels, very little fluctuation throughout the series was shown ($P = 0.96$).

In order to examine the relevance of reliability in visual field tests we analysed the 16 visual field series excluded from the study due to one or more tests in each series being 'unreliable'. The mean MD and 99% CI of these visual field series compared to the reliable series used in the main study ($n = 202$) is shown in Table 3. It is of note that although not quite statistically significant ($P = 0.08$),

Table 3 Mean MD and 99% CI of unreliable fields series

Level (AGIS)	Reliable, mean MD	Reliable, 99% CI	Unreliable, mean MD	Unreliable, 99% CI
Nil	-0.9	0.2	-1.5	0.6
Moderate	-9.7	0.6	-11.2	2.3
Severe	-16.9	0.8	-18.3	2.0

all levels of field defect available using the unreliable field series showed a 99% CI greater than that of the reliable series.

Discussion

Although our reported CI's should be viewed with caution when assessing individual visual field tests, we have shown that within a series of fields in a 'stable' eye, there is relatively little fluctuation in the MD. Although the size of the groups reduce as the field defects worsen, we believe that the sample sizes remain large enough for valid results, and show that the worse the defect, the greater the range of fluctuation, which has been suggested previously by Flammer *et al*,⁵ who only divided eyes into 'normal', 'suspect glaucoma', and 'glaucoma'.

Although the MD can be influenced by such conditions as cataract, our study should be free from such influences as all visual fields, based on first and last field analysis were stable over a long period of time. However, to ensure that initial visual field defects were not cataract induced, with the final defect being related to glaucomatous progression, the patients' ophthalmic records were examined. No patients included in the study had a cataract extraction during the period of the visual field series.

A potential criticism of our study is the variety of test strategies included in the study, however, our series represents a 'real-world' scenario often found in the clinical area. All tests analysed were Humphrey 24-2 visual fields with older visual fields being either Fastpac or Full Threshold, with more recent fields tending to be either Swedish Interactive Thresholding Algorithm (SITA)Fast or SITA Standard, although the majority of these recent fields were SITA Standard (Table 4). In a clinic situation these differing field tests may need to be compared. There are however theoretical differences between them. The SITA is designed as a faster test in comparison to the older Full Threshold strategies, and takes into account prior knowledge of the nature of visual field loss, reduces the need for false-positive catch trials, and speeds up the rate of stimulus presentation in patients who respond quickly.¹²

Table 4 Visual field strategies

Strategy	Number
Fastpac	459
Full threshold	75
SITAFast	44
SITA Standard	432
Total	1010

Table 5 Variation in MD between different test strategies

Level (AGIS)	Significant P-value
Nil	0.54
Mild	0.3
Moderate	0.21
Severe	0.2
End stage	0.11

Although some studies show that a certain amount of caution is required when comparing different test strategies,^{15–17} many studies concur that the SITA strategies are accurate, with the benefit of a reduced test duration over the Full Threshold strategies.^{18–22} In relation to the MD in studies which have compared different test strategies, a variety of results have been reported. Bengtsson *et al*²¹ showed that the MD did not differ significantly between the SITA Standard, SITAFast and Full Threshold strategies, as did Wild *et al*¹⁷ who included the FASTPAC strategy ($P = 0.291$). Budenz *et al*,¹⁵ however, showed that a statistically significant difference between the MD on the SITA Standard, SITAFast and Full Threshold strategies existed ($P < 0.001$). Sekhar *et al*¹⁶ showed that the repeatability of the MD was excellent except in the SITAFast strategy, and Budenz *et al*¹⁹ showed that when comparing SITA to Full Threshold a significant difference in MD was evident ($P < 0.001$). O'Brien *et al*²³ suggests a significant difference ($P = 0.007$) in the MD between Full Threshold and FASTPAC, although it is also argued that although FASTPAC may be less accurate than Full Threshold, this is offset by the benefits of a reduced test duration.

Using an ANOVA test with our data, there proved to be no significant variation between MD scores of the different test strategies at any level of field defect (Table 5), although the P -value does appear to reduce with an increase in the level of field defect, a factor which may require further investigation.

In our study, the potential for less accuracy in the faster tests becomes negligible as the main use of these strategies is to make the test more reliable for certain patients prone to fatigue, or as an initial test when a lack of experience is a potential concern. With the use of the

tests in this way, as well as there being no significant variation in the MD scores between the differing tests in our study, the results reported in our study remain valid.

This study has shown that in an eye with reliable and stable visual fields, one should expect only very little fluctuation in the MD, indicating that an increase in the MD beyond that of the 99% CI described may suggest progression, providing the ocular media remains clear. The reliability indices of the visual field test should therefore be regarded as a primary consideration when assessing a visual field for progression. When assessing the unreliable visual fields, we have shown a 99% CI greater than that of the reliable fields, however, this is worthy of further study as our sample of unreliable fields can be considered as too small. It is likely that the results of greatest value in a clinic setting are those which only incorporate the three visual fields in the centre of the series. That is, fields with no defect had a 99% CI of 0.3 dB, mild defects 0.4 dB, moderate defects 0.8 dB, severe 1 dB, and 1.3 dB for end-stage defects (Figure 1).

It is important not to make a false-positive diagnosis of progression in stable glaucoma as in prescribing unnecessary additional treatment there are inherent problems such as the increased risk of medication or surgical side effects, a likely reduction in quality of life and the cost to both the patient and the health-care provider. Multiple medication changes and polypharmacy which may result from a false-positive diagnosis of progression, may itself cause poor medication compliance^{24,25} and subsequent disease progression. Conversely, a worsening of a visual field MD within our CI may indicate real progression of the disease and this should be born in mind. However, as suggested by Hoskins *et al*⁴ and Schulzer²⁶ repeating a visual field test, especially if the reliability is in question may well confirm any suspicions without subjecting the patient to additional medication, surgery or indeed delaying needed treatment for too long.

Our study gives clinicians a guide to the confidence limits for the MD within which a new field should stay if it is to be considered stable by AGIS scoring, and therefore should be of value to those clinicians without computerised progression analysis facilities.

References

- 1 Heijl A, Asman P. Pitfalls of automated perimetry in glaucoma diagnosis. *Curr Opin Ophthalmol* 1995; **6**(2): 46–51.
- 2 Hutchings N, Wild J, Hussey M, Flanagan J, Trope G. The long-term fluctuation of the visual field in stable glaucoma. *Invest Ophthalmol Vis Sci* 2000; **41**(11): 3429–3436.
- 3 Stewart W, Kelly D, Hunt H. Long-term and short-term fluctuation in pattern discrimination perimetry. *Am J Ophthalmol* 1992; **114**(3): 302–306.

- 4 Hoskins H, Magee S, Drake M, Kidd M. Confidence intervals for change in automated visual fields. *Br J Ophthalmol* 1988; **72**(8): 591–597.
- 5 Flammer J, Drance S, Zulauf M. Short and long-term fluctuation in patients with glaucoma, normal controls and patients with suspected glaucoma. *Arch Ophthalmol* 1984; **102**: 704–706.
- 6 Boeglin R, Caprioli J, Zulauf M. Long-term fluctuation of the visual field in glaucoma. *Am J Ophthalmol* 1992; **113**(4): 396–400.
- 7 Hitchings R. *Fundamentals of Clinical Ophthalmology, Glaucoma*. BMJ: London, 2000, p 117.
- 8 Viswanathan A, Fitzke F, Hitchings R. Early detection of visual field progression in glaucoma: a comparison of PROGRESSOR and STATPAC 2. *Br J Ophthalmol* 1997; **81**: 1037–1042.
- 9 Gaasterland D, Ederer F, Sullivan K, Caprioli J, Cyrlin M. Advanced glaucoma intervention study 2 visual field test scoring and reliability. *Ophthalmology* 1994; **101**(8): 1445–1455.
- 10 Vingrys A, Demirel S. False-response monitoring during automated perimetry. *Optom Vis Sci* 1998; **75**(7): 513–517.
- 11 Haley M. *The Field Analyzer Primer*, 2nd ed. Allergan: California, 1987, p 134.
- 12 Henson D. *Visual Fields*. Butterworth Heineinan: Oxford, 2000, p 126.
- 13 Heijl A, Bengtsson B. The effect of perimetric experience in patients with glaucoma. *Arch Ophthalmol* 1996; **114**(1): 19–22.
- 14 Kulze J, Stewart W, Sutherland S. Factors associated with a learning effect in glaucoma patients using automated perimetry. *Acta Ophthalmol* 1990; **68**(6): 681–686.
- 15 Budenz D, Rhee P, Feuer W, McSoley J, Johnson C, Anderson D. Comparison of glaucomatous visual field defects using standard full threshold and Swedish interactive threshold algorithms. *Arch Ophthalmol* 2002; **120**(9): 1136–1141.
- 16 Sekhar G, Naduvilath T, Lakkai M, Jayakumar A, Pandi G, Mandal A *et al*. Sensitivity of Swedish interactive threshold algorithm compared with standard full threshold algorithm in Humphrey visual field testing. *Ophthalmology* 2000; **107**(7): 1303–1308.
- 17 Wild J, Pacey I, O'Neill E, Cunliffe I. The SITA perimetric threshold algorithms in glaucoma. *Invest Ophthalmol Vis Sci* 1999; **40**(9): 1998–2009.
- 18 Bengtsson B, Heijl A. Evaluation of a new perimetric threshold strategy, SITA, in patients with manifest and suspect glaucoma. *Acta Ophthalmol Scand* 1998; **76**(3): 268–272.
- 19 Budenz D, Rhee P, Feuer W, McSoley J, Johnson C, Anderson D. Sensitivity and specificity of the Swedish interactive threshold algorithm for glaucomatous visual field defects. *Ophthalmology* 2002; **109**(6): 1052–1058.
- 20 Wild J, Pacey I, Hancock S, Cunliffe I. Between-algorithm, between-individual differences in normal perimetric sensitivity: full threshold, FASTPAC, and SITA. *Invest Ophthalmol Vis Sci* 1999; **40**(6): 1152–1161.
- 21 Bengtsson B, Heijl A. Comparing significance and magnitude of glaucomatous visual field defects using the SITA and Full Threshold strategies. *Acta Ophthalmol Scand* 1999; **77**(2): 143–146.
- 22 Artes P, Iwase A, Ohno Y, Kitazawa Y, Chauhan B. Properties of perimetric threshold estimates from Full Threshold, SITA Standard, and SITA Fast strategies. *Invest Ophthalmol Vis Sci* 2002; **43**(8): 2654–2659.
- 23 O'Brien C, Poinoosawmy D, Wu J, Hitchings R. Evaluation of the Humphrey Fastpac threshold program in glaucoma. *Br J Ophthalmol* 1994; **78**(7): 513–514.
- 24 Gurwitz J, glynn R, Monane M, Everitt D, Gilden D, Smith N *et al*. Treatment for glaucoma: adherence by the elderly. *Am J Pub Health* 1993; **83**(5): 711–716.
- 25 Tsai J, McClure C, Ramos S, Schlundt D, Pichert J. Compliance barriers in glaucoma: a systematic classification. *J Glaucoma* 2003; **12**(5): 393–398.
- 26 Schulzer M. Errors in the diagnosis of visual field progression in normal-tension glaucoma. *Ophthalmology* 1994; **101**(9): 1589–1594.