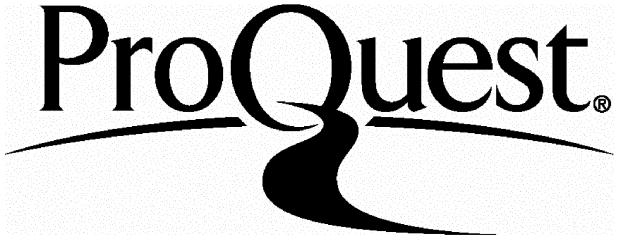
Visual Field Changes Following Trabeculectomy

Mark Wilkins
Institute of Ophthalmology and Moorfields Eye Hospital

Thesis submitted for the degree of
Doctor of Medicine
University of London
June 2001



Abstract

The use of automated visual fields to detect and monitor glaucoma is hampered by having no gold standard against which to compare them. In the case of monitoring disease progression visual fields display large amounts of fluctuation that can mask true change. The analysis of fields using pointwise linear regression (PLR) has been developed to more accurately detect change. However the criteria for change using PLR are themselves poorly understood. This thesis examines the collection of field data from a surgical trial of trabeculectomy and then explores the detection of change in the eyes in the study using conventional and PLR grading techniques.

Analysis of field data from an initial group of patients in the trial reveals the large amount of change detected using existing criteria. Much of the change detected is due to noise or fluctuations in a patient's response that do not represent real change. The use of modified criteria has variable effects on the detection of change. From this group of modified criteria, 6 can be selected on an empirical basis. All maximise the detection of progression while minimising improvement. Given the data available it is not possible to link any changes in visual field to changes in media opacity, especially cataract. When the selected criteria are tested against a) extended follow up data and b) a second group of patients from the same trial one criterion offers the ability to detect progression in both groups of patients while minimising the detection of improvement. This criterion requires a particular spatial arrangement of points in the field.

Analysing groups of patients' fields using PLR without regard to treatment offers a way of developing change criteria prior to analysis within treatment arms.

Table of contents

List of figures

List of tables

List of abbreviations

5FU	5-fluorouracil
asb	apostilb
AGIS	Advanced Glaucoma Intervention Study
ALT	argon laser trabeculoplasty
C	cortical
C-D	cup-disc
CIGTS	Collaborative Initial Glaucoma Treatment Study
CNTG	Collaborative Normal Tension Glaucoma
CPSD	corrected pattern standard deviation
DCCT	Diabetics Control and Complications Trial
dB	decibel
EMGT	Early Manifest Glaucoma Trial
ETDRS	Early Treatment of Diabetic Retinopathy Study
FovTh	foveal threshold
GH	general height
GHT	glaucoma hemifield test
GLASS	Glaucoma Screening Study
HFA	Humphrey field analyzer
IOL	intraocular lens
IOP	intraocular pressure
LED	light emitting diode
LOCS	lens opacity classification system
logMAR	logarithm of minimum angle of resolution
LONS	Longitudinal Optic Neuritis Study
LTG	low tension glaucoma

MD	mean deviation
MEH	Moorfields Eye Hospital
MMC	mitomycin c
mmHg	millimetres of mercury
MPS	Macular Photocoagulation Study
MRC	medical research council
NFL	nerve fibre layer
NC	nuclear colour
NO	nuclear opalescence
NIH	National Institutes of Health
NTG	normal tension glaucoma
ONTT	Optic Neuritis Treatment Trial
OHT	ocular hypertension
PERK	Prospective Evaluation of Radial Keratotomy
PLR	pointwise linear regression
PNFB	perimetric nerve fibre bundle cluster
POAG	primary open angle glaucoma
PSCLO	posterior subcapsular lens opacity
PSD	pattern standard deviation
RAFEP	regression analysis with fixed effects on panel data
RCOphth	Royal College of Ophthalmology
SF	short-term fluctuation
Sig	significance
SITA	Swedish interactive threshold algorithm
SLO	scanning laser ophthalmoscope
UCL	University College London
VA	visual acuity

Acknowledgements

I would like to thank the following for their help in producing this thesis:

Peng T Khaw for selecting me to work on the 5-FU trial and supervising this thesis.

Fred Fitzke also supervised this thesis and gave up time for numerous meetings about progression criteria.

Roger Hitchings provided support and constructive criticism.

Aachel Kotecha, Tom Lowe and Rachel Michell worked with me on the 5-FU trial, the data collected reflects their efforts as much as mine.

David Crabb for general support plus allowing me to use figure 11

Andromachi Frangouli who helped provide additional field data.

The MRC 5-FU trial is funded by the Medical research Council of Great Britain, grant number G 9330070.

Statement of originality

The work presented in this thesis was undertaken by myself at the Institute of Ophthalmology, London and at Moorfields Eye Hospital, London. The data was collected from an ongoing clinical trial based at Moorfields. My co-workers on the trial are acknowledged above. The thesis and data analysis represent my own work helped by the guidance of my supervisors. It has not been submitted for any other degree or award.

1 Introduction

1.1 Ocular Changes in Glaucoma

Glaucoma is a progressive optic neuropathy with characteristic structural damage to the optic nerve and characteristic visual field defects (Gupta and Weinreb 1997). A recent estimate put the number of people bilaterally blind from glaucoma by the year 2000 at 6.7 million (Quigley 1996). It is generally assumed that the glaucomatous disease process is modifiable by treatment. Attempts have therefore been made to reduce the impact of glaucoma by improved detection of cases and improved treatment. Hindering such efforts is the fact that no single measurable variable (such as blood pressure, or blood sugar) exists for diagnosing glaucoma and that there remains no ideal single method for following a patient once the condition has been diagnosed. Three broad areas are examined in clinical settings: intraocular pressure; optic disc and nerve fibre layer; and visual fields.

Although an abnormality in one area may be strongly indicative of glaucoma, it usually requires an abnormality in more than one area for a diagnosis to be made. The detection and monitoring of glaucoma with respect to these three areas is discussed in the next 3 sections.

1.1.1 Intraocular Pressure

Intraocular pressure (IOP) is no longer included in the definition of glaucoma. However it has been linked to it from the earliest description of the condition, where a hard painful eye was linked to blindness. Recently our views on the relationship between intraocular pressure and glaucoma have changed. Many epidemiological surveys have quantified the distribution of IOP in various populations and have established that the mean IOP is around 15-16 mmHg (Armaly 1965; Bankes et al. 1968; Kahn et al. 1977; Sommer et al. 1991; Klein et al. 1992; Bonomi et al. 1997). The mean can vary for different populations, being lower in Eskimos (13.6 mmHg in men tested using a Perkins

tonometer) (Arkell et al. 1987) and Mongolians (12.7 mmHg in men tested using a Goldmann tonometer) (Foster et al. 1996) than Caucasians (Armaly 1965; Bankes et al. 1968; Kahn et al. 1977; Sommer et al. 1991; Klein et al. 1992; Bonomi et al. 1997). Furthermore the distribution is not normal. Despite this IOP has frequently been analysed as though it were so (Colton and Ederer 1980). In particular the concept has developed that the upper normal limit of intraocular pressure is the mean + 2 standard deviations or approximately 21 mmHg. Given the fact that the distribution is not normal, it is not reasonable to use the figure of 21 as the cut off between normal (≤21 mmHg) and abnormal (>21 mmHg) in such a way.

Even though one cannot say that an IOP above a certain point is abnormal, raised intraocular pressure undoubtedly does play a significant role in the pathogenesis of glaucomatous damage. Unilateral conditions that produce raised IOP frequently lead to unilateral field loss. Several groups have retrospectively studied the association between intraocular pressure and visual field loss finding that increased intraocular pressure leads to an increased likelihood of visual field loss (Sommer 1989) and an accelerated rate of loss (Jay and Murdoch 1993). Other work has shown an association between optic disc changes and IOP in normal (Jonas et al. 1998) and high-tension glaucoma (Varma et al. 1995). A non systematic review by Palmberg showed an association between IOP and progression in glaucoma treatment studies (Palmberg 1996). Prospective epidemiological studies have shown an increasing risk of glaucoma with raised intraocular pressure (Armaly et al. 1980; Sommer et al. 1991; Dielemans et al. 1994; Leske et al. 1995; Leske et al. 2001). However at the same time the same studies have pointed out that 40-55% of their subjects with glaucoma had IOP's below 21 mmHg at screening (Sommer et al. 1991; Dielemans et al. 1994; Leske et al. 2001). The influence of factors other than IOP is shown in data from the Baltimore Eye Study (Sommer et al. 1991). The prevalence of glaucoma among black subjects was 4.3 times that of white subjects yet there was no difference in the mean IOP's between black and white subjects.

Despite a large number of treatment studies very little hard evidence exists to show that, in patients with glaucoma, lowering IOP reduces the rate of visual field progression. Two retrospective studies have shown some benefit from lowering IOP (Mao et al. 1991; Odberg 1993). In a systematic review of trials of glaucoma treatment Rossetti showed that glaucoma treatment is associated with a reduction in intraocular pressure (Rossetti et al. 1993). At the time that the paper was written only 3 randomised controlled trials with data on visual fields were available (Epstein et al. 1989; Kass et al. 1989; Schulzer et al. 1991). Their statistical combination failed to show a significant protective effect of active treatment. A subsequent prospective study randomising glaucoma patients to drops, argon laser trabeculoplasty (ALT) or trabeculectomy found that the differences in IOP between the ALT and trabeculectomy patients explained the differences in field progression between the 2 groups (Migdal et al. 1994). Despite similar mean IOP values in the laser and medical groups the greater deterioration of visual fields in the medical group allowed the authors to conclude that "medical treatment appears to make the fields worse, or allows them to deteriorate faster, in some way beyond its effect on the IOP". The Glaucoma Laser Trial randomised both eyes of patients with glaucoma to initially receive argon laser trabeculoplasty in one eye and timoptol 0.5% in the other (Glaucoma Laser Trial Research Group 1991). Additional treatment was prescribed according to a pre-defined list of additional topical treatments. Eyes initially receiving ALT had a consistently lower IOP than those initially receiving timoptol. Automated field test results from the initial study when the bulk of the patients were still under follow up showed a trend for improvement in the fields in both arms of the trial. It was only with extended follow up after 4 years that quantitative and qualitative grading of fields showed more fields deteriorating than improving in either arm. Unfortunately at 4 years the follow-up rate was only 120/271 patients (Glaucoma Laser Trial Research Group 1995). At this point the mean point sensitivity in the ALT group was 0.7 dB higher than at enrolment, whereas the figure in the medical treatment group was unchanged. In contrast subjective grading of the optic discs

was more negative, both arms had a majority of patients with deteriorating discs. There was no significant difference in the changes between the two arms. The researchers were thus only able to claim that initial treatment with ALT was at least as efficacious as initial treatment with topical medication. Despite the lower IOP achieved using ALT they were unable to make any claims linking IOP and automated visual field changes. In the Advanced Glaucoma Intervention Study the patients were randomised to different sequences of treatment using ALT and trabeculectomy (AGIS 1994). Claiming "unexpected findings of statistically significant interactions" the authors published their results separately for black and white patients (AGIS 1998). Having done this they felt that better visual field preservation only occurred in white patients who achieved better IOP reduction. The lack of concordance between IOP lowering and field preservation in black patients was they felt due to other factors. The only discussion of potential factors is a reference to a review that cites age, race, myopia, and vascular haemodynamics as having a role in the development and progression of glaucomatous visual fields (Drance 1997).

Airaksinen followed the optic discs from normal, ocular hypertensive and glaucoma subjects for 5 years using planimetry to quantify any changes (Airaksinen et al. 1992). None of the normals showed any change in rim area while 79% of the glaucoma subjects and 43% of the ocular hypertensives did. Further analysis showed that 90% of the change in rim area loss over time was accounted for by variables other than the ones measured in the study namely age, disc area, initial rim area and IOP. There was a weak correlation between rate of rim area change and IOP ($R^2 = 0.058$, $P<0.05$) that was weaker than that seen with age ($R^2 = 0.09$, $P<0.01$).

There is little evidence linking the role of IOP in the conversion of eyes with ocular hypertension to glaucoma. Prospective (Kitazawa et al. 1977) and retrospective (David et al. 1977; Odberg 1993) studies of patients with ocular hypertension have failed to find an association between IOP and those patients who develop field loss. In a prospective study of timoptol treatment versus placebo for ocular hypertension, the placebo group had a

greater rate of field conversion than the timoptol group (Kass et al. 1989). The mean IOP in the timoptol group was lower than in the placebo group. No statistical comparison of IOP's for all patients converting versus those who did not was given. Finally in a retrospective study of trabeculectomy patients IOP could not predict which patients would show deterioration of visual field (Kidd and M 1985).

The Collaborative Normal-Tension Glaucoma Study Group randomised patients with normal tension glaucoma to no treatment or to treatment designed to lower their IOP by 30% (Collaborative Normal-Tension Glaucoma Study Group 1998; Collaborative Normal-Tension Glaucoma Study Group 1998). The treated arm experienced less visual field progression. However the only link between IOP and field progression was obtained when data regarding patients developing cataract was censored.

The lack of conclusive evidence between IOP and glaucoma progression has led to the establishment of the Early Manifest Glaucoma Trial (EMGT) to evaluate the effectiveness of reducing IOP in early, previously untreated open-angle glaucoma. Patients with IOP's <30 mmHg and with mild field loss have been randomised to either treatment or no treatment at all. Both arms will be closely monitored (Leske et al. 1999).

Intraocular pressure remains a risk factor but is not the sole mechanism by which glaucomatous damage occurs. Measuring IOP has some value in detecting patients with high pressures who are at increased risk of developing glaucoma. Similarly measuring IOP has some role in monitoring patients undergoing treatment. Other than at extremely high levels IOP cannot be exclusively relied upon to diagnose or monitor glaucoma.

1.1.2 Optic Disc

The normal and glaucomatous nerve fibre layer and optic disc have been studied with a view to developing strategies to detect and monitor glaucomatous damage. Both qualitative and quantitative techniques have been used.

Crucial to using the optic disc to detect glaucoma is the ability to decide what constitutes a normal disc. The size and variability of normal and glaucomatous discs in black and white populations has been defined using a population based study (Varma et al. 1994). Normal discs are variable in size with significant racial and sexual variations. In blacks the optic disc is approximately 12% larger than in whites. Blacks also show a larger interindividual variation in their disc area varying from 0.9 to 6.28mm^2 while whites vary from 1.15 to 4.94 mm^2 (Varma et al. 1994). Male optic discs are on average 2-3% larger than female optic discs (Varma et al. 1994).

Further study of disc components reveals further variability. The neuroretinal rim area is proportional to disc size in blacks and whites (Britton et al. 1987; Caprioli and Miller 1987; Varma et al. 1994). However the limits of normality for rim area cover a very large range from 0.7 mm^2 to more than 4 mm^2. The regression slope of rim area against disc area is smaller for blacks than whites (Varma et al. 1994) suggesting that as black optic discs increase in size there is a smaller increase in rim area compared to white optic discs. For any given disc size the rim area is lower in blacks than in whites (Varma et al. 1994). It has also been shown that larger optic discs have a greater number of nerve fibres in humans (Jonas et al. 1992) and monkeys (Quigley et al. 1991). Since rim area increases with disc size it is likely that rim area is an indirect marker for the number of nerve fibres in an optic disc. Thus it has been suggested that for any given disc size a black optic disc will contain fewer nerve fibres than a white one (Varma et al. 1994). In glaucoma it is the death of ganglion cells that produces visual loss (Wygnanski et al. 1995). Initial reports of photoreceptor death (Panda and Jonas 1992) in glaucoma have not been confirmed in subsequent reports (Kendell et al. 1995; Wygnanski et al. 1995). The earliest visible changes due to ganglion cell death in glaucoma are defects in the retinal nerve fibre layer (Hoyt and Newman 1972). The changes may be seen as a diffuse thinning or as wedge shaped defects in the retinal nerve fibre layer. These changes may precede optic disc and visual field changes (Tuulonen and Airaksinen 1991; Tuulonen et al. 1993), they may

themselves be preceded by an optic disc haemorrhage (Airaksinen et al. 1981). With progressive fibre loss the NFL becomes thinner and/or wedge shaped defects enlarge.

When sufficient ganglion cells are lost, optic disc morphology changes. The changes seen can be variable. Diffuse neural rim thinning, vertical elongation of the cup, rim notch formation, or pallor of the rim with no change in outline have all been described (Airaksinen et al. 1992). The loss of ganglion cells also leads to alterations in the position of blood vessels, the loss of rim tissue removes support for the vessels. Changes in the optic disc may precede development of visual field defects (Sommer et al. 1979; Motolko and Drance 1981; Funk 1991; Zeyen and Caprioli 1993; Kamal et al. 1999).

With progressive ganglion cell death there is an increase in cup size and in the cup-disc (C-D) ratio. Cup size and C-D ratio are both a function of the size of the optic disc and rim area, which as mentioned above, vary quite considerably in the normal population. Inter observer agreement in detecting C-D ratio is poor (Lichter 1976; Varma et al. 1989; Varma et al. 1992).

Another feature of human glaucomatous eyes is an outward bowing of the lamina cribrosa (Quigley et al. 1983). This has also been observed in an animal model of glaucoma (Coleman et al. 1991). It is believed that bowing can cause kinking of axons, interrupt axonal transport and thus cause cell death. Reduced connective tissue support in the superior and inferior quadrants of the optic disc is believed to contribute to the preferential loss of axons at these sites (Quigley and Addicks 1981; Miller and Quigley 1988). Such preferential loss produces vertical enlargement of the optic cup. Where the rate of loss of ganglion cells is asymmetric between the eyes the optic cup size will also be asymmetric.

The accelerated loss of ganglion cells in glaucoma leads to a reduction in rim area that is greater than the natural decline seen with age (Airaksinen et al. 1992). However it has not been possible to discriminate between normal and glaucomatous eyes using rim area because of the overlap between the two groups (Caprioli 1992; Damms and Dannheim 1993).

Several studies have looked at the sensitivity and specificity of techniques for separating glaucomatous from normal optic discs. Features assessed include vertical C-D ratio (Damms and Dannheim 1993), photographs of nerve fibre layer defects (Airaksinen et al. 1984), circumlinear vessel baring (Balazsi and Werner 1983), and scanning laser ophthalmoscope (SLO) images (Wollstein et al. 1998). When assessing the impact of these studies it is important to look at the patient selection. Using features in disc photographs from a screened population in the Baltimore Eye Study shows that there is no cut-off point for vertical C-D ratio or narrowest rim width when separating normal optic disc from glaucomatous ones (Tielsch et al. 1991). Using quantitative techniques, such as planimetry and digitised image analysis of videographic images, to predict field loss from optic disc features the best sensitivity obtainable was 74% (O' Connor et al. 1993). This figure was achieved by obtaining nerve fibre height measurements; poor specificity reduced the diagnostic precision value to 68%. (Diagnostic precision = the total proportion of eyes that were correctly identified as having healthy or glaucomatous eyes). The same article found that qualitative assessments based on optic disc and nerve fibre photographs had higher levels of diagnostic precision. By analysing SLO images, high sensitivity and specificity at separating glaucomatous disc from normal ones have been obtained. Using the 99% prediction interval from the linear regression between the optic disc area and the log of the neuroretinal rim area 96.3% specificity and 84.3% sensitivity were produced (Wollstein et al. 1998). With this technique we may be closer to an objective technique for separating early glaucomatous eyes from normal ones. However this analysis technique has not been tested on its ability to monitor disease progression, nor has it yet been used in a large population study, being derived from a hospital population.

Serial examinations of the optic disc and nerve fibre layer have tried to establish the temporal link between optic disc and visual field changes. In a study of 813 ocular hypertensive eyes examined annually over 5 years Quigley et al were able to identify 37 eyes that developed field loss and 37 age matched controls that did not (Quigley et al. 1992).

The authors compared optic disc and nerve fibre layer photographs in the 2 subgroups. Disc change was detected in only 19% of converters, while progressive nerve fibre layer atrophy was observed in 49% of converters. A smaller study followed one eye from 15 patients over 6 years (Zeyen and Caprioli 1993). Eight eyes with an initially normal visual field showed disc changes using planimetry; six of these eyes did not develop field abnormalities. Population studies have also highlighted the mismatch between optic disc and visual field changes. In the Beaver Dam Eye Study of the 104 cases of "definite" POAG 45 had high IOP, abnormal visual fields, but normal stereo disc photographs (Klein et al. 1992).

Segmental analysis of the optic disc using either planimetry (Weber et al. 1990) or more frequently the scanning laser ophthalmoscope (Asawaphureekorn et al. 1996; Anton et al. 1997; Yamagishi et al. 1997; Anton et al. 1998)has shown a correlation between focal disc changes and regional visual field loss. However problems remain with considerable interindividual variability in optic nerve head size and configuration; some field zones topographically map to certain optic disc rim areas with greater predictability than others (Yamagishi et al. 1997; Anton et al. 1998).

1.1.3 Visual Field

The visual field of an eye refers to all the space that can be seen at any given instant. The visual field is quantified by presenting light stimuli and determining whether the eye can see them or not. It requires that not only is an image formed within the eye but information about the image must be transmitted to the brain and then "perceived". The stimuli used to map a visual field may be static or moving, and can have variable size and intensity. The stimuli may be projected on to a flat surface so that more peripheral stimuli are further from the eye (Campimetry), or they may be projected onto a curved surface or bowl, thus keeping the stimulus-eye distance constant. An isoptre is an imaginary line joining those points in the visual field with the same visual threshold. If the separation

between isoptres is known then a 3 dimensional model of sensitivity can be constructed. Although not the first person to map the visual field it was Traquair who coined the phrase "hill, or island, of vision" to describe the decrease in sensitivity to light as one moves peripherally.

It has been suggested that early glaucoma produces a generalised contraction of isoptres/ reduction in sensitivity (Anctil and Anderson 1984; Caprioli et al. 1987; Drance 1991). Other work has tended to contradict this view suggesting that early signs of glaucoma in the visual field are represented by regional depressions in sensitivity or scotomas (Werner and Drance 1977). Indeed specific attempts to detect fields with diffuse loss purely due to glaucoma have been unsuccessful (Werner et al. 1982; Heijl 1989; Åsman and Heijl 1994). Åsman and Heijl using the glaucoma hemifield test found that only 2 out of 1582 eyes showed diffuse field loss (Åsman and Heijl 1994). Some of the discrepancies between papers over their detection of diffuse loss have been attributable to methodology: static vs. kinetic perimetry, the confounding effect of miosis and cataract, patient selection and the technique used to detect depression. Global indices (Caprioli et al. 1987) have been used to indicate diffuse depression rather than specific algorithms (Åsman and Heijl 1994). Recent work that took into account the presence of lens opacity saw the incidence of diffuse loss in early glaucoma fall from 12.4% to 4.4% once patients with cataract were excluded (Chauhan et al. 1997). Henson, using patients with established early field loss, has argued that diffuse loss is present as a component of the total field loss (Henson et al. 1999). The results of his study differ from those of Heijl's (Heijl 1989) even though both use the same technique measuring the sensitivity of the 10 best points in the field. However Heijl's conclusions were based on eyes that may well have had even earlier field loss, indeed 2 of them had no field loss at all.

It is now recognised that isolated defects represent the more common initial field defect in glaucoma (Drance 1969). The position of these can be paracentral, nasal or temporal (Werner and Drance 1977; Hart and Becker 1982). Prior to the development of a

scotoma the visual field displays an increase in short term fluctuation (Flammer et al. 1984) in the threshold sensitivity to light.

Progression of glaucoma is manifest by 1) scotomas becoming deeper 2) enlargement of scotomas and 3) development of new scotomas (Mikelberg and Drance 1984). Despite our knowledge of glaucomatous field progression the ability of humans to agree on which field series are progressing is poor (Werner et al. 1988). Furthermore algorithms to select deteriorating fields from stable ones are also not good enough for the algorithm to be relied upon alone (Smith et al. 1996).

The loss of visual field represents a late sign of damage in glaucoma. In post-mortem studies Quigley has shown that 25-50% of neurons die before detection of field loss using manual perimetry. With automated perimetry a loss of 20% of ganglion cells correlates with approximately a 5dB loss of sensitivity and a 40% loss of ganglion cells correlates with approximately a 10 dB reduction. These studies by their very nature are retrospective and small in size making comparisons difficult. The automated perimetry study consists of only 6 eyes, all taken from patients age 70 or older (Quigley et al. 1989). The manual perimetry study consists of 26 eyes, with donors being as young as 46 (Quigley et al. 1982).

Early perimetry concentrated on using white stimuli on white backgrounds. In an attempt to detect visual loss earlier in the glaucomatous disease process, other testing modalities have been tried. In glaucoma there is initially a greater loss of large ganglion cells. The loss occurs in humans (Quigley et al. 1987; Quigley et al. 1988) and in monkeys (Glovinsky et al. 1991; Glovinsky et al. 1993; Wygnanski et al. 1995; Desatnik et al. 1996) and occurs centrally as well as peripherally (Wygnanski et al. 1995). The preferential loss of larger neurons has stimulated interest because it is known from rhesus monkey studies that the optic nerve contains principally 2 types of neuron. Ninety percent are the smaller diameter slower conducting P cells, the remaining 10% percent are mainly the larger diameter faster conducting M-cells. M ganglion cells have larger receptive fields, higher

contrast gain, and are believed to be involved in motion detection and to preferentially receive inputs from cones sensitive to shorter wavelengths. The preferential loss of M cells in glaucoma has led to field testing using short-wave perimetry (Johnson et al. 1993; Johnson et al. 1993), motion detection (Fitzke et al. 1989), contrast sensitivity (Arden and Jacobson 1978), and flicker (Lachenmayr et al. 1991; Lachenmayr et al. 1991; Yoshiyama and Johnson 1997). Recent work has challenged the idea that test modalities that are selective for a particular type of cell death will offer any benefits in the detection of glaucoma (Johnson et al. 2000). Shrinkage in ganglion cell size prior to death may be causing an artefactual loss in large ganglion cells. At present white on white perimetry remains the dominant field testing technique in the United Kingdom. Table 1 shows the type of field test used on patients included in the Royal College of Ophthalmologists trabeculectomy audit (Beth Edmunds – personal communication).

A major problem with visual field testing are the intertest fluctuations at individual points that make detection of change over successive fields difficult (Heijl et al. 1987; Heijl et al. 1989). Fluctuation between field tests can lead to a false impression of change, with fields appearing to improve as well as progress. As will be discussed later a variety of data acquisition and analysis techniques have been developed to improve the detection of glaucoma and to improvement the monitoring of the disease.

No technique has yet been shown to have overall superiority in detecting and monitoring glaucoma. In the future, sophisticated optic disc analysis may offer this. At the present time visual fields still have a role in managing patients with glaucoma.

Table 1 Distribution of field test techniques according to the RCOphth Trabeculectomy Audit.

Field Test	%
Humphrey	56.3
Friedman	12.6
Goldmann	11.8
Henson	10.4
Octopus	3.6
Dicon	2.1
Tangent Screen	1.5
Other	1.8

Automated perimetry (Humphrey, Henson, Octopus and Dicon) is used on over 70% of the patients in the audit; while kinetic perimetry (Goldmann and tangent screen) is used to test less than 15%.

1.2 Automated perimetry for measuring visual fields

The first formal attempts at measuring the visual field were performed by Young in 1801. Fifty years later in 1856 von Graefe published his accounts of campimetry, the plotting of visual fields on a flat surface. He mapped the blind spot, scotomas, hemianopsias, and described isopter contraction. The following decade Förster, using an Arc perimeter, extended the area tested to beyond 45 degrees. In the 1950's Goldmann developed his hemispheric projection perimeter that is still in use today. He also quantified the relationship between the area and luminance of a test object.

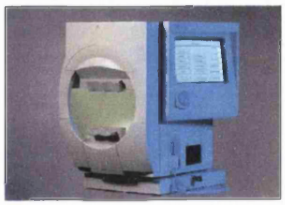
Early visual field work was based around the concept of kinetic perimetry; the subject indicates when a moving stimulus can first be seen. In 1939 Sloan described the use of static perimetry wherein the stimulus is not moved but its intensity is varied. Harms and Aulhorn went on to design the Tübinger perimeter that permitted static and kinetic perimetry. Subsequently the Armaly screen was developed. It uses a combination of kinetic and static manual perimetry on a Goldmann perimeter to screen a patient for glaucomatous field defects. At most points stimuli of variable size and intensity are presented successively until a spot is seen. The blind spot and 2 temporal parts of the field

are mapped using standard kinetic techniques (Rock et al. 1971; Rock et al. 1973). This technique produced fast test times with high sensitivity and specificity figures for separating patients with glaucomatous field loss from normals. The test is not widely used but the point locations used have been incorporated into screening programs available in the Humphrey field analyser. Although attempted by others in the early 1960's, Lynn and Tate were the first to demonstrate an automated static perimeter in 1969. The rapid change in technology over the next 30 years has led to the development of modern automated perimeters the most popular of which is the Humphrey Field Analyzer (table 1).

1.2.1 Humphrey Field Analyzer

The Humphrey Field Analyzer (Allergan-Humphrey, San Leandro, California, USA) is a projection automatic perimeter. Models I and II consist of a single unit containing a projection bowl, a display screen for entering information, a printer and data storage facilities in the form of a floppy disk (I and II) or hard drive (II). Visual field testing is performed with the patient facing the stimulus bowl, a white hemispherical bowl with a radius of 330 mm. Two diffuse light sources are used to illuminate the bowl so that the background luminance is 31.5 asb. The illumination is checked when the machine is switched on and at the beginning of each test with the patient seated in front of the bowl. In addition the local background luminance is tested before each stimulus is presented so that stimulus intensity may be adjusted to cope with any local variations in background luminance. Spot stimuli lasting 0.2s are projected on to the surface of the bowl using a mirror. Step motors control the position of the mirror. Using neutral density filters, stimulus strength can be varied from 0.08 to 10,000 asb (51 dB). Stimulus size can be varied to match the 5 sizes (V to I) available on the Goldman perimeter. Coloured filters are supplied that allow colour perimetry such as "blue on yellow" to be performed.

Figure 1 Humphrey Field Analyzer Mk II



The Humphrey offers screening and threshold strategies (see below) that can be deployed over a variety of test patterns. Threshold test patterns cover the central 10, 24 or 30 degrees of the central visual field or the peripheral field from 30 to 60 degrees. The points in these tests are laid out in a grid pattern. Other threshold tests are available to assess specific areas of peripheral field, neurological or macular function. A choice of 2 test patterns is available to test the central 24 and 30 degrees and the peripheral 30-60 degrees (xx-1 and xx-2). For each area covered the patterns use the same point spacing (6 degrees for the central tests, 12 degrees for the peripheral tests) however the positioning of the points is different being offset by 3 degrees horizontally and vertically in the case of the 24- and 30- tests. It is possible to combine -1 and -2 tests to obtain an even more detailed visual field.

1.2.2 Octopus

The Octopus Perimeter (Interzeag AG, Schlieren, Zurich, Switzerland) is a projection automatic perimeter that shares any common features with Humphrey Field Analyzer. In its present form it is a single computer driven unit with a 330mm diameter projection bowl. The Octopus uses a 0.1s stimulus duration with a 4 apostilb background intensity. Colour perimetry is also possible. Threshold and screening test patterns are available. Several of the threshold test patterns are the same as those on the Humphrey Field Analyzer. Patterns 31 and 32 correspond to the 30-1 and 30-2 respectively.

Figure 2 Octopus 101 and 1-2-3 perimeters

The 101 model is shown on the left, the 1-2-3 model on the right

1.2.3 Other Perimeters

The Dicon range of perimeters (Coopervision, California, USA) use an illuminated bowl in which light emitting diodes (LED's) have been mounted. The background illumination is white with a standard luminance of 31.5 asb. The LED's produce light with a peak emission of 570 nm, which is in the yellow-green region of the visible electromagnetic spectrum. The LED's are arranged along radials with eccentricity increments of 2.5 degrees within the central 10 degrees, of 5 degrees within the 10 to 30 degree circles, and of 10 degrees peripheral to that. Screening and threshold testing strategies are available.

Figure 3 Dicon LD400 perimeter

1.3 *Automated field test strategies*

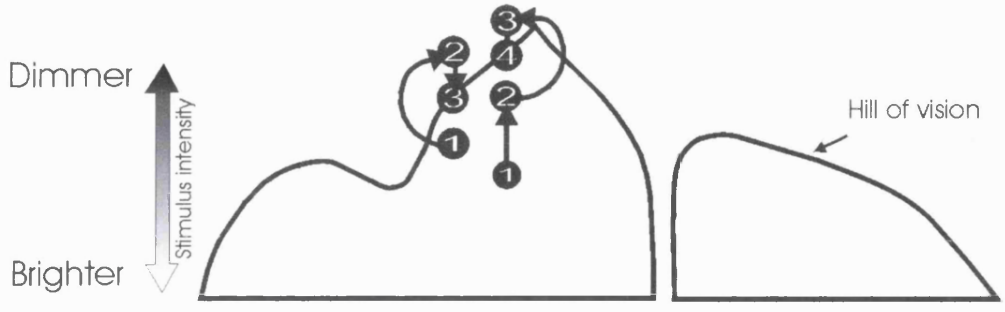
1.3.1 Thresholding

The threshold for any given point in a visual field is not a fixed stimulus intensity but is more accurately defined by a frequency of seeing curve. It is frequently taken to be the stimulus intensity that is seen 50% of the time. Eliciting a full frequency of seeing

curve at each test point would be too time consuming. Rather threshold values are assumed to lie between the closest seen and unseen stimuli.

Automatic perimeters typically use a staircase strategy to elicit threshold values. A stimulus is presented that is close to expected threshold. If it is seen successive stimuli are made progressively weaker until they are not seen. The process is then reversed so that stimuli are made more intense until they are seen again. The threshold for that point is either the average of the last seen and unseen points or taken as the first seen value after reversal. If the initial stimulus is not seen then the whole process is reversed. Both the Humphrey Field Analyzer and the Octopus use a 4-2 strategy when testing threshold. Initial changes in stimulus strength are in 4 dB steps until threshold is crossed. After reversal stimulus strength changes in 2 dB steps (Figure 4). The Humphrey uses the last seen stimulus as its value for threshold. The Octopus uses the average of the last seen and not seen stimuli for its value. Values that are 5 dB outside the expected value are retested using the staircase method again. The result of the second test is displayed in brackets under the initial threshold value on the printout (Figure 7). In deciding the strength of the initial stimulus the Octopus uses an age corrected value selected from a database. The Humphrey thresholds 4 primary points at the start of the test, one in each quadrant. It then uses the values from these tests as the basis for the surrounding secondary points. Values from these secondary points are then used to calculate the initial stimulus strength for surrounding points.

Figure 4 Full threshold staircase strategy as used the Humphrey Field Analyzer.



Initial changes in stimulus strength are in 4 dB steps until threshold is crossed. After reversal stimulus strength changes in 2 dB steps (Based on fig 18 "The Field Analyzer Primer", Humphrey Instruments Inc)

In ascribing a numerical value to the threshold determined on the Humphrey, the perimeters maximal intensity stimulus is assigned a value of 0 dB. A 20 dB stimulus is thus 2 log-units less intense than the maximum stimulus. Although the Humphrey Field Analyzer can present dimmer stimuli it is generally thought that most humans cannot respond to stimuli below 40 dB i.e. 1/10,000 of the maximum stimulus. Thus the range of the machine is from 0-40 dB (Anderson and Patella 1999).

An additional option speeds up the threshold strategy by using threshold values from the last test as the starting point. Staircase strategies have been compared using a computer model and results from normals (Johnson et al. 1992). Both reveal that the present 4-2 strategy offers the best trade-off between efficiency and accuracy.

1.3.2 Screening

Screening strategies are designed to quickly detect subjects with visual field defects. In separating abnormal from normal fields it is important that the screening test has acceptably high sensitivity and specificity. Unfortunately there is frequently a trade off between sensitivity and specificity. The simplest form of screening uses a stimulus that is expected to be slightly above threshold. The choice of stimulus intensity may be the same for the whole area tested (one-level testing) or it may be eccentricity-compensated. The second option is more efficient given the natural decrease in retinal sensitivity with increasing eccentricity from the fovea. With one level testing, depending on the stimulus level chosen relative defects may be missed or false positives generated.

There is a large interindividual variation in retinal sensitivity in subjects with normal fields. Some factors such as age have a predictable effect on threshold whereas others such as pupil size and media opacity are not predictable. To compensate for this many strategies will actually test the threshold of early points to determine the general level of the field under test. Once a defect is found on a screening test it may simply be recorded as a

missed point or it may be tested further. Additional stimuli can be used to fully threshold the missed point or to classify it as a relative or total defect (threshold related screening). The Humphrey Field Analyzer offers single level screening, plus 3 threshold related strategies. Having tested 4 primary points the threshold related strategies use stimuli 6 dB above the expected level. The "Supra Threshold" strategy retests all missed points, recording them as seen or not seen depending on the second presentation. The "Three Zone" strategy tests missed points at maximum stimulus intensity. If the point is seen then it is recorded as a relative defect if it is still not seen then it is recorded as an absolute defect. The "Quantify Defects" strategy performs a full threshold test at all missed points. The Octopus screening strategy uses a 3-zone strategy to classify points as "normal", "relative defect", or "absolute defect".

1.3.3 Reliability Indices

In HFA screening and threshold strategies several indices are generated as a guide to the reliability of the subjects responses:

False positive – The test subject hears the stimulus projector move as if about to show a target at a new location. The projector does not produce a stimulus. A trigger happy patient will respond to the noise of the projector moving generating a false positive response.

False negative – supra threshold stimuli are represented at points where threshold has already been determined. If the patient fails to respond then a false negative response is recorded.

Fixation loss – At the beginning of the test the blind spot is mapped out. Later on stimuli are represented to the blind spot at random intervals. If the patient has not lost fixation these stimuli should not be seen.

For all these indices the number of errors is expressed as a percentage of maximum number possible. Using the criteria suggested by the manufacturer (<33% false-positive,

<33% false-negative, <20% fixation losses) a high percentage of patients are found to be unreliable when initially tested using automated perimetry on a Humphrey Field Analyzer. As many as 45% of patients with glaucoma, 35% of ocular hypertensives and even 30% of normals have been found to be unreliable (Katz and Sommer 1988; Bickler-Bluth et al. 1989). False negative responses are significantly more common in patients with glaucoma compared to ocular hypertensives or normals (Heijl et al. 1986; Katz and Sommer 1988; Katz et al. 1991). Poor fixation produces less general depression, and reduced localised defects in patients with glaucoma (Katz and Sommer 1990). High false positive rates reduce the depression seen in glaucomatous fields as well (Katz and Sommer 1990). Serial visual field testing, with a test interval of approximately one year has not been shown to reduce reliability scores (Katz et al. 1991).

1.3.4 Swedish Interactive Threshold Algorithms - SITA

Developed for the HFA II, the Swedish Interactive Threshold Algorithms have recently been developed in an attempt to obtain quicker threshold measurements while retaining an accuracy comparable to the existing test (Bengtsson et al. 1997).

At the beginning of the test the HFA creates an internal mathematical model. The model consists of 2 probability curves for each point. The curves describe the probabilities of threshold values for that point if it were normal or abnormal. Each curve is derived from data from normal and abnormal fields. The data utilised includes for each point: a) age-corrected threshold values b) frequency of seeing curves c) correlations between the threshold of a point and the values at other locations.

Testing commences as in the standard full threshold test with a full threshold test of one point in each of the 4 quadrants. Threshold values from these points are used in turn to select initial stimulus intensities at adjacent points. Testing then occurs at adjacent points using staircase testing procedures. After each stimulus presentation the probability of threshold curves for the tested and adjacent points are recalculated. Periodically all

probability of threshold curves are recalculated. Testing at a point is stopped when a predetermined level of threshold certainty is reached. However at least one reversal of threshold must have occurred before testing of a point has been reached.

Additional algorithms in the software estimate the false positive and negative responses rates, and maximize the rate of presentation of stimuli. False positives are estimated from responses that occur during periods when no response is anticipated, thus no additional testing is required. False negatives are estimated from the pattern of responses along with traditional catch trials.

Testing of normals (Bengtsson et al. 1998), ocular hypertensives (Bengtsson and Heijl 1998), and subjects with glaucoma (Bengtsson and Heijl 1998) has confirmed a large reduction in test time of approximately 50%. In normals SITA has 1.9 dB higher sensitivity than full threshold test, in ocular hypertensives and glaucoma subjects the figure was 2.4 dB. This difference is attributed to a reduction in fatigue due to shorter test duration. In all groups there was a similar test-retest variability compared to the full threshold test.

1.3.5 Variability in threshold values during perimetry

A problem with static perimetry is the fluctuation in a patient's response when shown stimuli at the same location. This fluctuation can be seen if the same point is thresholded twice within the same test (short-term fluctuation – SF) or if the thresholds from 2 separate tests are compared (long-term fluctuation). Each point in the visual field has its own frequency of seeing curve. The curve is altered by the disease process.

1.3.5.1 Global Measures of Variability

Various equations have been developed to generate a numerical index of short-term fluctuation based on testing the threshold at a given number of points more than once. A generalised equation (Flammer et al. 1985) takes the square root of the pooled variance:

Where x_{ij} is the j^{th} threshold test at each point i, m is the number of points tested and n is

$$SF = \sqrt{\frac{\sum_{i=1}^{m}\sum_{j=1}^{n}(x_{ij} - x_{i\bullet})^2}{m(n-1)}}$$

the number of times each point is tested. For the Humphrey threshold testing strategy m=10 and n=2 which should yield the equation:

$$SF = \sqrt{\frac{\sum_{i=1}^{m}(x_{i1} - x_{i2})^2}{m}}$$

However the actual equation used is similar to:

$$SF = \sqrt{\frac{\sum_{i=1}^{m}(x_{i1} - x_{i2})^2}{2m}}$$

Giving values that are consistently $\sqrt{2}$ larger. Factors influencing the estimate of SF have been modelled using a computer simulation of a patient undergoing a visual field test (Casson et al. 1990). Using a computer program named KRAKEN it is possible to test perimetric strategies against a software "module" that contains representations of visual fields. The software has been programmed to respond to the testing in a similar manner to a human subject, there are modifiable response characteristics such as reaction time, fluctuation, fatigue and errors. The number of points tested (m), how many times each point is tested (n), as well as the variability of the response can all be varied. The underlying testing strategy is the same as that of the Humphrey staircase threshold strategy. Results from the program show that for an equal number of threshold tests, the standard deviation of SF falls as the number of locations falls. Testing 5 locations 4 times produces less variation in SF than testing 10 locations twice. However although the Humphrey testing strategy tests 10 locations twice it weights the contribution of each point using the normal intra-test variance of each point. Further comparisons using this model are thus impossible. It is important not to have too few points tested since an unrepresentative fraction may lie in an area of field loss producing a distorted value of SF.

When considering intertest variability it is possible to use all points tested to generate global indices of variability. Werner et al suggested 2 indices 1) the square root of the mean variance of all test locations for a subject and 2) mean of the range of all test locations (Werner et al. 1989). The 2 indices were calculated using a group of 20 patients with stable glaucoma who had at least 4 Octopus fields (program 32 – most points tested once) over a year. The average total variability per subject is 2.8 dB using the variance-based calculation and 5.1 dB using the range-based calculation (Werner et al. 1989). Using range as the measurement of variability, 95% of points had a variability of less than 13 dB.

Program G1 on the Octopus tests each location twice. Using the mean of individual pointwise variances (Boeglin et al. 1992) comparable figures to Werner's for global variability were found. In a retrospective analysis of fields patients deemed to have stable glaucoma had a mean variance of 7.0 dB^2 while patients with unstable glaucoma had a variance of 9.7 dB^2 ($P<0.0005$). Range of variability as defined above increased as the initial sensitivity declined. There is no correlation with of long-term fluctuation with eccentricity once after correcting for the change in initial sensitivity associated with eccentricity.

1.3.5.2 Cluster Measures of Variability

If the 74 points tested using program 32 on the Octopus are divided in 10 clusters and the mean threshold of each cluster calculated then cluster variability is 7 dB (Werner et al. 1989). It still requires a minimum of 2 clusters to change by 7 dB for the probability of this being a random event to fall to below 5%.

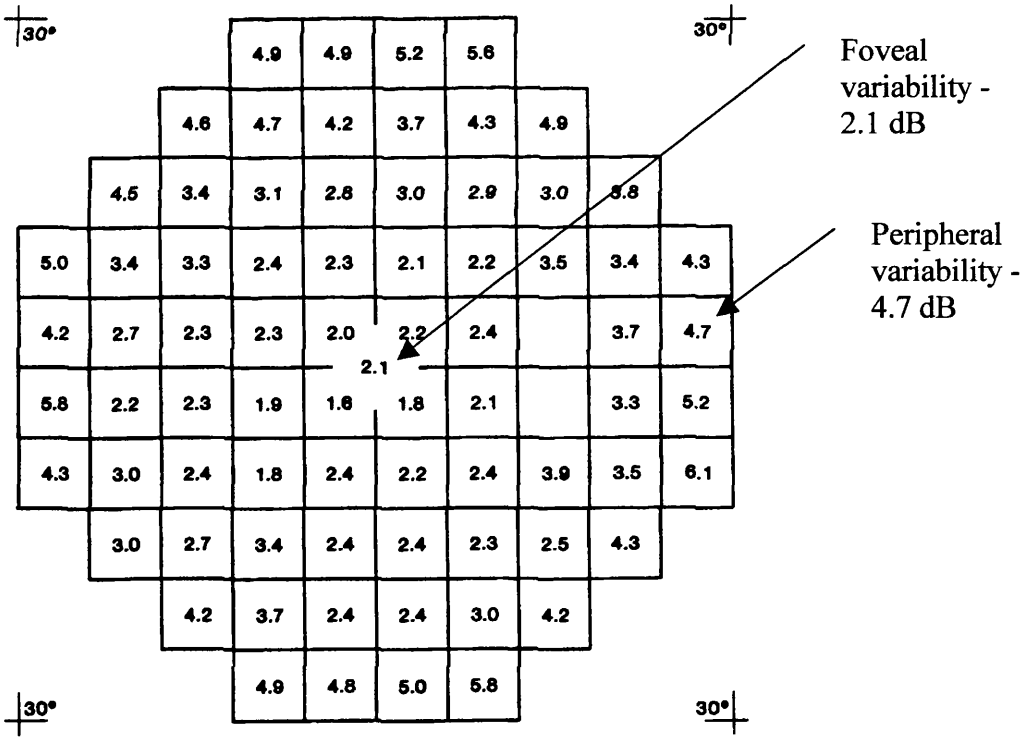
The Advanced Glaucoma Intervention Study (AGIS) developed a scoring system (see below) for Humphrey visual fields based on the number and depth of clusters of adjacent depressed test sites in the upper and lower hemifields and the nasal area (AGIS 1994). The study required patients to have a repeat visual field within 60 days of the first qualifying field. Scores from the initial field test ranged from 1-17. A large inter test fluctuation was deemed to be a change in score of 4 or more. Using this criterion for

fluctuation, the score on the second field test was found to improve in 11% of patients and worsen in 5%. The likelihood of a large change in field score increased if the second field test was performed more than 1 week after the first compared to 1 week or less. Intertest fluctuation in field score was not related to the patient's age.

1.3.5.3 Pointwise Measures of Variability

Normals: The variability of the visual field in normal patients has been characterised in a large study where subjects were randomly selected from a computerised population database (Heijl et al. 1987). Using rigorous exclusion criteria, field data (Humphrey 30-2) from 74 normals tested on 3 separate visits was obtained. Only data from the second and third visits was used. From this data the intraindividual intertest variation and the interindividual variation were calculated (Figure 5).

Figure 5 Intraindividual intertest variation – in dB.



Sensitivity variability in normals, at the same point, between tests, more than doubles as you move from the foveal area to the periphery of a 30-2 field (Heijl et al. 1987)

The data shows that threshold variability generally increases with distance from fixation. This applies when comparing tests a) repeated in the same normal individual and b) between normals. Furthermore the distribution of threshold at each point is non-Gaussian. Variability also increases with distance from fixation within a single test. One possible explanation for the effect of eccentricity on threshold variability is test artefact from lens holders, ptosis and prominent brows. However the effect begins too centrally for this to be the main reason. The authors suggest that the effect is due to a reduction in "neural channel" density as you move into the periphery leading to a reduced signal-to-noise ratio. A smaller study corroborated this on a variety of perimeters (Lewis et al. 1986). Variability also increases at the edge of scotomas (Haefliger and Flammer 1989; Haefliger and Flammer 1991) whether they are physiological (blind spot) or pathological (glaucoma), the effect being greater for the latter (Haefliger and Flammer 1989).

A specific testing algorithm to quantify response variability compared variability at 4 locations in eyes with optic neuritis, ocular hypertension, glaucoma as well as normals. Variability increased as sensitivity decreased, furthermore the relationship between the variability and sensitivity was similar for all 4 subject groups. This suggested that regardless of the pathological mechanism for ganglion cell loss it is the remaining ganglion cell density that determines sensitivity and variability (Henson et al. 2000).

Ocular hypertensives: A retrospective study looked at the fields of controls and patients diagnosed as ocular hypertensives on the basis of repeated normal manual screening static perimetry and IOP > 21 mmHg (Werner et al. 1982). When compared with age and sex matched controls the ocular hypertensive patients had a greater variability of threshold responses at the 10 points tested in the central visual field. However the subjects with raised IOP had larger cup to disc ratios than the controls suggesting that optic nerve damage had occurred in some of them. It is known that prior to the development of a scotoma the visual field displays an increase in short term fluctuation (Flammer et al. 1984) in the threshold sensitivity to light.

Glaucoma Subjects: Studies have looked at intertest fluctuation in the fields of glaucomatous patients, with prospective studies of multiple fields over short periods of time (Heijl et al. 1989) and retrospective long term studies of patients deemed to be clinically stable (Werner et al. 1989; Boeglin et al. 1992). Heijl et al (Heijl et al. 1989) tested one eye of 51 patients with glaucoma once a week for 4 weeks. All field tests were conducted using the 30-2 program on the HFA; subjects were not naïve to automatic perimetry. The range of field loss ranged from incipient to advanced. Points were ranked according to their deviation from expected normal threshold. Using the percentile distribution of subsequent thresholds at each point it is shown that points with initially moderate loss showed marked variation in threshold, ranging from normal values to absolute defect. Variability is much lower in points showing initially normal or near normal sensitivities. Having divided the test area into concentric zones they showed that variability increased with eccentricity but only for those points whose initial sensitivity is 10 dB or more below normal. Averaging the first 2 test tests reduced subsequent inter test variability. A retrospective study looked at patients with stable and unstable fields tested using Program G1 on the Octopus (Boeglin et al. 1992). A minimum of 3 fields tested over one year plus previous automated perimetric experience was required. Their pointwise results for stable patients were similar to those of Heijl showing a decrease in variability as the initial point value approaches normal. Variability was greater with increasing distance from fixation and was also greater in the superior hemifield compared to the inferior hemifield. However after correcting for the decreased sensitivity associated with greater eccentricity there was no longer a correlation between eccentricity and inter test variability. There was no correlation between fluctuation and age. Both of the above papers (Heijl et al. 1989; Boeglin et al. 1992) point out the great deal of variability seen in depressed points. Testing on the HFA shows that for points depressed by –8 to –18 dB on initial testing the 95% prediction interval for subsequent tests cover the range from normal sensitivity to absolute defect. Similarly testing glaucoma patients using the Octopus

showed that for points with a sensitivity of 15 dB or less the 90% prediction interval covers almost zero to normal values (Boeglin et al. 1992). Given Werner's variability figure of 13 dB for each point and the 74 points in the program 32 field he calculates there is almost a 90% chance of observing a point randomly fluctuating by 13 dB or more (Werner et al. 1989). Only when 7 points show a change of 13 dB or more is the probability of this being a random event <5%. This has very significant implications in trying to detect glaucomatous visual field progression.

1.3.6 FASTPAC

Comparisons of FASTPAC and full threshold screening in normal (Flanagan et al. 1993) and glaucomatous (Flanagan et al. 1993) patients showed that fields performed with FASTPAC have a greater short-term fluctuation. Furthermore this effect is more marked with increasing age.

1.4 The impact of test conditions on visual fields

1.4.1 Patient Age

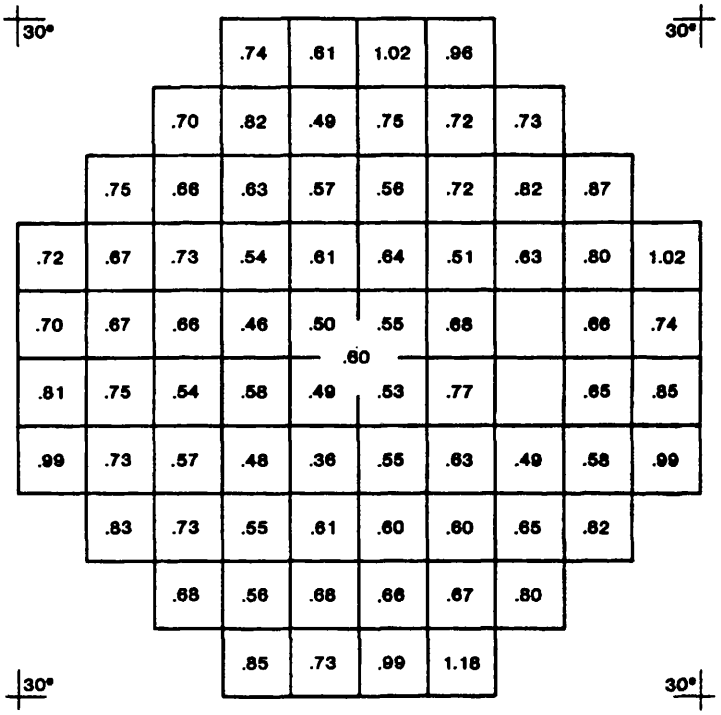
Studies using kinetic as well as automated perimetry have tried to quantify the effect of age threshold values on normals. None of these studies represent a longitudinal follow up of normals but rather rely on comparing normals of different ages. Kinetic perimetry suggested (Drance et al. 1967) that mean sensitivity decreased by 1 dB per decade. Two studies with automated perimetry have confirmed the finding and allowed more detailed analysis.

Figure 6 shows the pointwise age slopes (loss of sensitivity in decibels per decade) in normal patients undergoing a 30-2 test on the HFA (Heijl et al. 1987). Sensitivity decreases with age at all points the change being more marked in the periphery.

An earlier study using the Octopus and a linear regression analysis showed a mean sensitivity fall of 0.58 dB per decade over the whole field with greater fall in sensitivity in the periphery (Haas et al. 1986). Furthermore the fall was greater in the superior hemifield compared to the inferior one. Attempts to explain the change in sensitivity with age have postulated that opaque ocular media, reduced pupil size, reduced photopigment

absorbance, and cellular death at the retinal and post-retinal level could be responsible. Johnson et al found similar age related decreases in sensitivity in normal patients undergoing 1) yellow on yellow perimetry to minimise the effect of an aging lens 2) yellow on yellow perimetry with a large bright stimulus to minimise the effect of a small pupil and 3) normal perimetry (Johnson et al. 1989). They felt that age related changes were not due to pre-retinal factors but due to changes at the level of the retina.

Figure 6 Loss of sensitivity in dB per decade across the central 30°.



Pointwise rates of loss of sensitivity vary across the field from 0.36-1.18 dB/decade. The rate decreases from fovea to mid periphery before increasing as one moves into the periphery (Heijl et al. 1987)

1.4.2 Refraction

Normals tested on automated perimeters while wearing deliberate spectacle over corrections show reduced retinal sensitivity (Weinreb and Perlman 1986). Over corrections in the range +1.0 to +2.0 D produce average decreases in sensitivity of 1.26-1.4 dB/dioptre. However the effect of over correction on retinal sensitivity is far less than might be suspected by their effect on visual acuity. A +2.0 lens reduces logMAR visual acuity from 0.0 to 0.33 (Heuer et al. 1987).

1.4.3 Pupil size

By pharmacologically altering pupil size Wood and colleagues showed that perimetric sensitivity increased with pupil size (Wood et al. 1988), the effect being greater for peripheral than central locations. Conversely it has been shown that pilocarpine produces a reduction in sensitivity in normals (Lindenmuth et al. 1989). Reversing pilocarpine induced miosis in glaucoma patients produced an improvement in mean defect (mean change 3.14 dB) with the improvement in threshold being more marked in the periphery (Rebolleda et al. 1992). Finally glaucoma patients given pilocarpine show a deterioration in mean defect, the deterioration is more marked the greater the initial mean defect is. Furthermore a greater deterioration is seen if the pupil is miosed to a diameter of 2 mm or less (Webster et al. 1993).

1.4.4 Lens Opacity

Although there are occasional reports of glaucomatous field defects being caused by cataract (Phillips et al. 1978; Thomas et al. 1996) it is now known that cataract usually causes a diffuse reduction in sensitivity in the visual fields of normals (Guthauser and Flammer 1988; Lam et al. 1991) and glaucoma patients (Lam et al. 1991; Budenz et al. 1993). The degradation of vision from cataracts comes from absorption, scatter, and blur. By using neutral density and diffusing filters to simulate cataract it has been shown that small media changes with negligible effects on visual acuity may have significant impact on automated retinal threshold sensitivities (Heuer et al. 1988). The effects of neutral density and diffusing filters have a far greater influence on perimetric threshold measurements than might be expected from their effect on visual acuity. This is the reverse of the situation with spectacle over correction where a small over correction produces significant but small change in threshold at the expense of large change in visual acuity (Heuer et al. 1986). Simulating cataract in glaucoma patients using a diffusing filter produces similar diffuse

depression of retinal sensitivity. The depression is of equal magnitude in scotomatous and "normal" areas (Budenz et al. 1993).

Field testing before and after cataract surgery in normals (Guthauser and Flammer 1988; Lam et al. 1991) and glaucoma patients (Stewart et al. 1995; Chen and Budenz 1998) (Smith et al. 1997) has also helped characterise the effect of media opacity on visual fields. Two studies on normals have suggested that threshold recovery following cataract extraction may be greater outside the central visual field. Guthauser (Guthauser and Flammer 1988) found that there was a negative improvement correlated with eccentricity while Lam (Lam et al. 1991) found that only the outer points of a Humphrey 30-2 test showed reduced improvement compared to the other points. Objective grading of cataract using digital densitometer (Guthauser and Flammer 1988) revealed a stronger correlation between lens opacity and mean density changes than between opacity and preoperative visual acuity. Subjective clinical grading of cataract was a poor predictor of postoperative threshold changes (Lam et al. 1991). Preoperative VA only significantly correlated with changes in foveal threshold and not with other thresholds. One retrospective review of glaucoma cases undergoing cataract extraction with or without a combined trabeculectomy found large changes in MD and foveal threshold (Chen and Budenz 1998). Furthermore preoperative foveal threshold was significantly associated with post operative changes in mean deviation and foveal threshold following surgery. Preoperative visual acuity was not associated with postoperative changes in mean deviation. However an earlier study found a small mean improvement of only 1.68 dB (Smith et al. 1997). A small study failed to find any change in field indices in patients undergoing phacotrabeculectomy (Stewart et al. 1995).

Because cataract induced changes in visual fields are an important potential confounder the AGIS study tried to assess the effect of cataract on visual field and acuity (AGIS 2000). They reported an improvement in visual field and visual acuity defect scores.

Their original conclusions on the effect of race-treatment interactions persisted albeit at a lower level. Their conclusions were limited by the lack of formal assessment of cataract.

The change in the shape of the "hill of vision" induced by cataracts also depends on the type of lens opacity and the perimetric technique. With non-nuclear cataracts threshold was raised to a greater extent at an eccentricity of 30 degrees compared with fixation when measured with projected size III stimuli on an Octopus perimeter, the reverse was true when sensitivity was measured with the small LED stimuli of a Dicon perimeter. Conversely, nuclear cataracts increased retinal threshold to a greater extent at the fovea compared with more peripheral regions for both the large projected and small LED stimuli (Wood et al. 1989).

The pattern standard deviation (PSD) plot of the Statpac printouts for Humphrey perimeters displays the significance of each points deviation from age-normal values after correcting for any generalised decrease or increase in media opacity (discussed above). Theoretically they should not affected by the diffuse change in sensitivity caused by cataract (Bengtsson et al. 1997). However at least one study has found a significant increase in the mean CPSD of subjects with dense scotomas following cataract extraction (Smith et al. 1997).

1.4.5 Supervision

A recent prospective trial randomised patients to supervision versus no supervision during perimetry (Van Coevorden et al. 1999). In the unsupervised arm of the trial patients a technician was present only for the first 30-60 seconds of the test. The study found that there might be a small potential benefit from supervision but that overall global indices were not significantly different between the 2 groups. The authors concluded that only those patients at risk of low reliability need to be supervised for the duration of the test.

1.4.6 Learning

Learning is the improvement in sensitivity, not attributable to long or short term fluctuation, seen when visual field tests are repeated. It is particularly common in patients previously naïve to perimetry. Initially described in manual perimetry it has been extensively studied with automated perimetry.

An early retrospective study looked at patients with previous manual perimetric experience who were tested using automated perimetry (Werner et al. 1988). No learning effect, as defined by a significant change in mean sensitivity, was detected. There was however a significant fall in short term fluctuation moving from field one to field two.

In contrast a prospective intensive study of 10 young normal males undergoing 8 Octopus field tests showed that volume sensitivity (linearly related to mean sensitivity) increased over the first 5 daily tests (Wood et al. 1987). This effect persisted on days 15,16 and 44 when the subjects were retested. The learning effect was greatest in the superior field and for eccentricities greater than 30 degrees. A subsequent larger prospective study of normals on the HFA confirmed i) the initial improvement in sensitivity quantifying it as 1-2 dB over the first few tests and ii) the learning effect was largest for peripheral points (Heijl et al. 1989). The study was sufficiently large to show that a small minority of patients showed a large improvement in sensitivity with repeat testing. The patients that did so tended to have concentric contraction of their visual field with low sensitivities of their peripheral points.

In a subsequent prospective study, 20 glaucoma suspects underwent daily bilateral (right before left) 30-2 fields for 3 days followed by a final a further test 12 days later (Wild et al. 1989). Seventeen patients completed the follow up. Global, central, peripheral, superior and inferior mean sensitivity each significantly increased from the first to the second right field tests and from the third to the fourth left field tests. At the same time short-term fluctuation, central mean defect and number of stimulus presentations showed a significant decrease between the first and second tests in the right eye. No further

significant change was noted with later tests in the right eye or for any tests in the left eye. The lack of improvement seen in the left eye was interpreted as being due to either a transfer of learning from right to left eyes and/ or a fatigue effect since the left eye was tested second. Global sensitivity was significantly higher in the left eye on the first test compared to the right eye. Most indices in the right eye showed no significant change beyond test 2 suggesting a plateau effect. It should be noted that this study contained a heterogeneous mixture of patients. By the time that the study had finished 3 diagnoses were possible in the patients under test: POAG, OHT and LTG. The study demonstrates the phenomenon of learning but its applicability to individual patient types remains limited. When the same patients were followed up with 2 tests on 2 successive days approximately 9 months later the learning effect initially observed had disappeared (Wild et al. 1991).

A more recent study used glaucoma patients unfamiliar to perimetry at entry who were subsequently shown to have field defects (Heijl and Bengtsson 1996). Mean deviation improved from the first to the second test but did not significantly change thereafter over tests 2 to 5. Learning was again shown to increase with eccentricity. Fields with moderate field loss showed greater learning than those with severe or mild loss. Short term fluctuation did not change with experience.

Recognition of the effects of learning has led to suggestions that the initial field be removed from any analysis (Flammer et al. 1984) or that 2 or 3 baseline fields be performed within a short period of time (Hoskins et al. 1988).

1.5 *Detecting Glaucomatous Field Loss*

1.5.1 Observer Opinion

To my knowledge there are no studies of the ability of human observers to separate normal and glaucomatous automated fields. With Goldmann fields grading systems have been used to quantify degree of field loss using either the area of remaining visual field (Sponsel et al. 1983; Smith 1986) or how close the field defect comes to threatening

fixation (Jay and Allan 1989). Only one of these techniques has undergone any validation which established the degree inter- and intra-observer agreement in a group of known glaucomatous fields (King et al. 1997). No study has looked at the ability of observers with or without the help of a grading system to separate normal from glaucomatous fields.

1.5.2 Global Indices

Global indices are used to numerically summarise an automated visual field. One of the simplest global indices devised uses the sum of all threshold values (Holmin and Krakau 1980). This index has only been used to follow patients with glaucoma and not to try and diagnose glaucomatous field loss. More sophisticated indices are generated by the STATPAC program for the HFA have been used to try and separate glaucomatous from normal fields. In an extensive analysis using a large database of normal visual fields Heijl et al developed the Statpac software available with the Humphrey Field Analyzer (Heijl et al. 1986). It is developed to "facilitate interpretation of single fields and to illustrate changes over time in consecutive threshold fields". The model in the software incorporates the following concepts:

1) The sensitivity at each point decreases linearly with age. The rate of decrease differs for different points in the field. The height and shape of the field of vision thus changes with time.
2) The normal inter-individual and intra-individual inter-test variability is not constant but varies with location.
3) The deviations of points from normal reference values are not assumed to follow Gaussian distributions but are derived using empirical data. The field database used to derive them contains multiple tests from many of its subjects. One test was randomly selected and the pointwise deviations of the other field tests calculated. This process was then repeated 50 times enabling percentiles of pointwise deviations to be calculated.

4) The model is based on normal fields from patients with previous perimetric experience. This reduces the limits of normality.

5) The model is based on normal fields with reasonable reliability indices (<20% fixation losses, <33% false negatives, and <33% false positives). Again this reduces the limits of normality.

Data from a field test is displayed numerically and graphically in the test printout (Figure 7). The absolute threshold values are displayed in map form at the top accompanied by a greyscale representation. Below this the total deviation and the pattern deviation are displayed as numerical maps and probability maps. The total deviation numerical plot is the difference between the actual threshold and the expected age adjusted threshold. The pattern deviation numerical plot emphasises the shape of the measured field. The measured field is elevated or depressed towards the level of the normal reference field before the measured field is pointwise subtracted from the normal one. In the probability maps the deviation of the measured field from the normal field is compared with the prediction limits for normality. The significances are displayed as greyscale maps. Four global indices are calculated (Heijl et al. 1986):

Mean Deviation (MD), a weighted deviation from the normal reference field. The value estimates the uniform part of the deviation.

$$MD = \left\{ \frac{1}{n} \sum_{i=1}^{n} \frac{(x_i - N_i)}{S_{1i}^2} \right\} / \left\{ \frac{1}{n} \sum_{i=1}^{n} \frac{1}{S_{1i}^2} \right\}$$

x_i is the measured threshold at point i

S_{1i}^2 is the variance of normal field measurements at point i

N_i is the normal reference threshold at point i

n is the number of test points

MD becomes more negative as a field defect progresses or if there is a generalised depression of field sensitivity such as with a cataract.

Pattern Standard Deviation (PSD), a weighted standard deviation of the point-wise differences between the measured and normal reference fields. PSD estimates the non-uniform part of any measured field deviation, this is frequently interpreted as a measure of the shape of the hill of vision. A small PSD indicates little deviation between the measured hill of vision and the normal reference field. A large value is consistent with localised field defects.

Short-Term Fluctuation (SF), the weighted mean of the standard deviations of the 10 points tested twice in a 24-2 or 30-2 field.

Corrected Pattern Standard Deviation (CPSD) estimates that part of non-uniform deviation which is not caused by SF. It is interpreted as the true deviation of the hill of vision from normal.

$CPSD^2 = PSD^2 - k*SF^2$

Where the constant $K>1$. The authors have not released their method for calculating K.

1.5.2.1 The Collaborative Initial Glaucoma Study Scoring System

The Collaborative Initial Glaucoma Treatment Study (CIGTS) scoring system generates a number to indicate the severity of field loss (Katz 1999). The score is generated from the HFA's total deviation plot (Figure 7). Locations are deemed to be depressed if the probability value of the measured threshold is ≤ 5%. A score is given to each depressed point if it has 2 neighbouring depressed locations. The score depends on the probability values of the 3 contiguous depressed points (table 2).

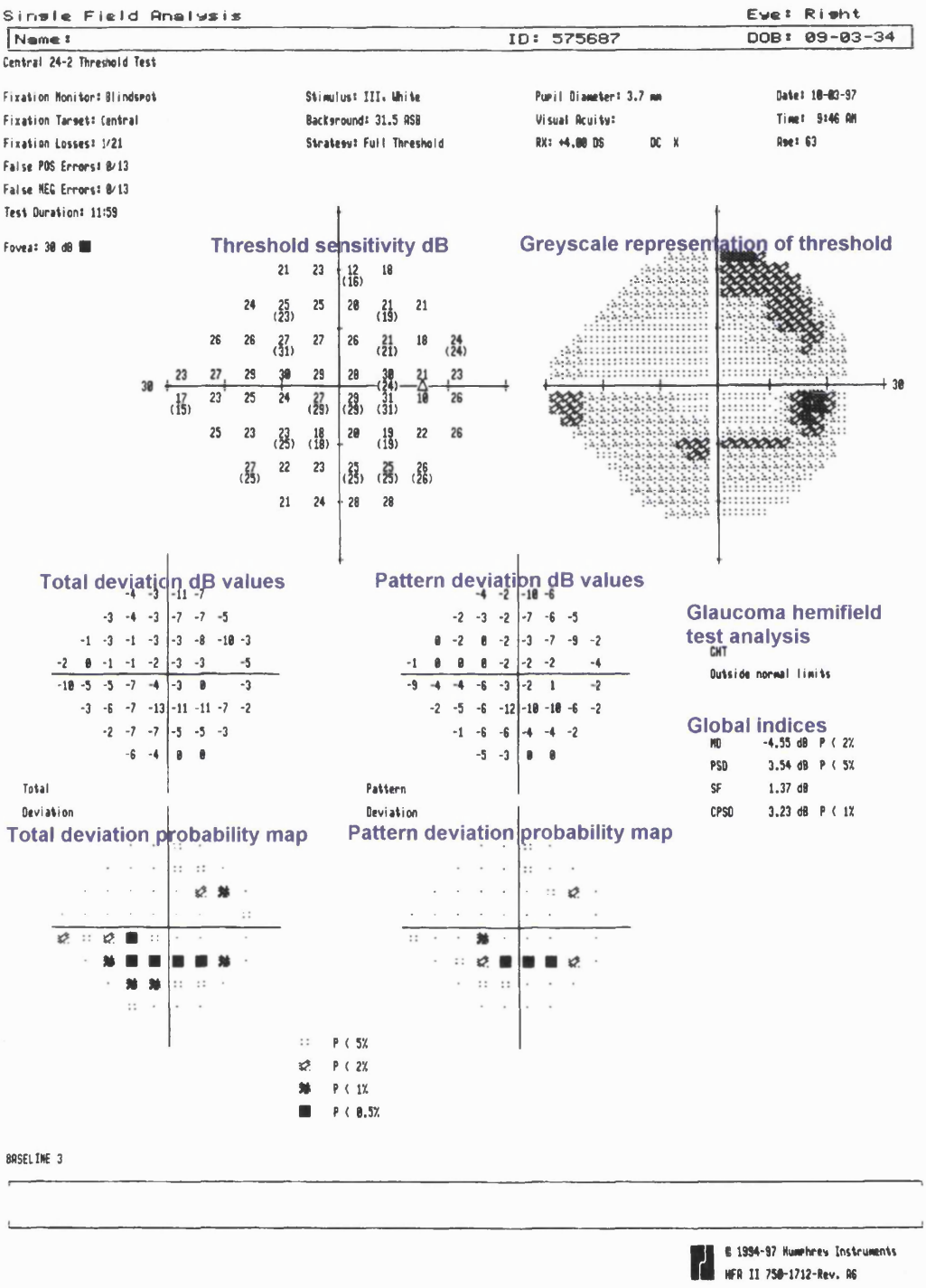
Table 2 CIGTS scoring system.

Probability Value	Score
5%	1
2%	2
1%	3
0.5%	4

Points awarded to each point if the point and at least 2 of its depressed neighbours are depressed by the amounts shown (Katz 1999).

The total score for all 52 points (2 points above and below the blind spot are omitted) is divided by 10.4 to produce a final score ranging from 0 to 20.

Figure 7 The Statpac printout from the Humphrey Field Analyzer



At the top are the threshold plot and its greyscale representation. Below it are the total deviation and pattern deviation dB plots and their probability plots. Total deviation=difference between threshold sensitivity and age-corrected normal sensitivity. Pattern deviation=total deviation minus General Height an index reflecting overall field depression and long term fluctuation. The probability plots highlight locations where the deviation exceeds that found in fewer than 5%, 2%, 1%, or 0.5% of normals.

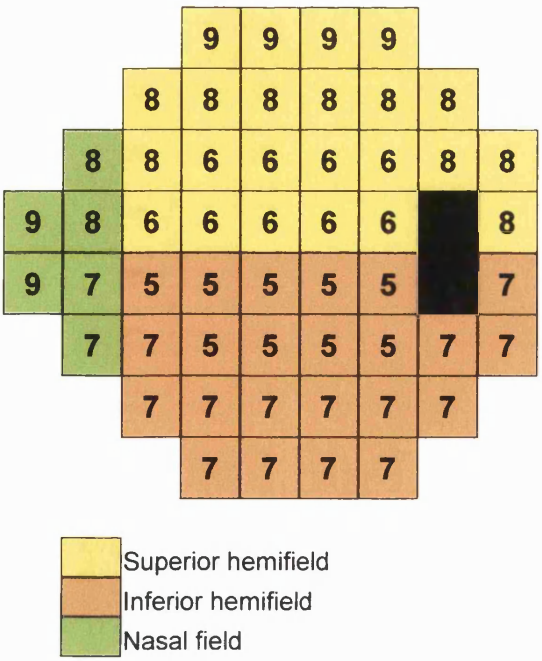
1.5.3 Cluster Analysis

This refers to the numerical summarising of sectors of the visual field. In general it is designed to pick up focal changes in the field that might be lost in global indices.

1.5.3.1 AGIS scoring

The Advanced Glaucoma Intervention Study (AGIS 1994) developed a scoring system that grades the test reliability and defect severity of 24-2 fields performed on a HFA. The grade of defect severity can be used to define whether a field is glaucomatous or not. Furthermore it can be used in longitudinal studies to assess whether there has been any field progression. The field defect score is calculated using values for each test point that represent the depression from the age-related normal database stored inside the HFA. These values are shown on the total deviation plot on the STATPAC-2 printout (Figure 7). Test sites above and below the blind spot are excluded. The actual values used and the layout of the 3 clusters are shown in Figure 8.

Figure 8 Cluster arrangement for the AGIS scoring system



Minimum amount of depression (dB) required to label a point as defective. Each coloured area represents a cluster of points. The field defect score for each cluster is calculated by assessing the number and arrangement of depressed points. The total field score is the sum of the 3 cluster scores.

Note
1 that the superior and inferior hemifields respect the horizontal midline while the nasal field does not
2 the minimum defect scores are not symmetrical across the horizontal midline

The amount of depression required to render a point "defective" is location dependant and is close to the value found in no more than 5% of age-matched normals. Furthermore the field is divided into 3 regions nasal, upper hemifield, and lower hemifield. Defects in the hemifields are only considered if 3 or more adjacent points in a single hemifield are depressed. Points are adjacent if they touch either side by side or obliquely. More than one cluster may occur in a hemifield. A cluster of 3 or more points in the nasal area constitutes a nasal defect. The cluster may cross the horizontal midline c.f. clusters in the hemifields, which may not. A single depressed point in the nasal area constitutes a nasal step. The field score is calculated using the following process:

1. For a nasal defect or step, add one to the score. If 4 of the 6 nasal points are depressed, add one more to the score.
2. In each hemifield with one or more clusters; add the relevant points to the score according to how many defective sites are in the clusters:

No. Sites in clusters	No. Points added to score
3 – 5	1
6 – 12	2
13 – 20	3
>20	4

3. If half or more of the adjacent defective sites in a hemifield are depressed by the amount shown add the relevant points to the score.

Amount ½ the points depressed by (dB)	Points added to score
28 or more	5
24 or more	4
20 or more	3
16 or more	2
12 or more	1

4. If a hemifield lacks a cluster of 3 depressed sites but contains 2 adjacent points depressed by 12 dB or more add 1 to the score.
5. Points for the hemifields and nasal area are summed

Table 3 AGIS Visual Field Reliability Rating For HFA test 24-2

CRITERIA	RATING
1. Total number of questions asked	
≤ 400	0
>400	1
2. Fixation Losses	
<20 trials	0
≥20 trials	
<20% fixation losses	0
≥20% fixation losses	1
3. False-positive responses	
<8 trials	0
≥8 trials	
<33% false-positive responses	0
≥33% false-positive responses	1
4. False-negative responses	
<8 trials	0
≥8 trials	
<33% false-negative responses	0
≥33% false-negative responses	1
5.Short-term fluctuation (dB)	
≤4.0	0
>4.0 but ≤6.0	1
>6.0 but ≤7.0	2
>7.0	3

An overall reliability score is generated by adding the scores from each of the above 5 areas. Fields scoring 3 or more are considered unreliable (AGIS 1994)

The reliability index is calculated using table 3. A field test is considered reliable if it scores 0, 1, or 2. Fields with a reliability score of 3 are deemed unreliable. The authors point out that this is an arbitrary division of field tests into reliable and unreliable, not based on any evidence.

The recent report from the AGIS study used a definition of an increase in visual field score of 4 or more to define a "Decrease of Visual Field", pointing out that this happens spontaneously in less than 5% of cases (AGIS 1994). The scoring system has also been compared with other algorithms in their ability to detect whether a single field can be scored as glaucomatous or normal (Katz et al. 1991). Other workers have used the AGIS scoring system to demonstrate the severity of field loss (Kamal et al. 1999).

1.5.3.2 Collaborative Normal Tension Glaucoma Study

The collaborative Normal-tension Glaucoma (cNTG) study used its' own definition of glaucomatous field loss and of glaucomatous progression (Schulzer 1994). Entry to the study required that 3 points on a Humphrey 24-2 or Octopus program 32 be depressed 5 dB from the normal value for age, with at least one point being depressed by at least 10 dB. All 3 points had to be on one side of the horizontal meridian. Such a defect had to be confirmed on 2 out of 3 fields performed within 4 weeks. The initial criterion for progression was a decline of at least 10 dB or 3x the average baseline short-term fluctuation by 2 adjacent points within or adjacent to a baseline defect. Furthermore the sensitivity of the declining points must be outside the range of all values for that point on the 3 baseline examinations. Once progression was detected it needed to be confirmed on repeat testing. The patient was given two attempts 1-4 weeks following the suspect test. From initial analysis of the results it became clear that excessive levels of progression were being detected (Schulzer et al. 1991; Schulzer 1994). Therefore if progression was detected on the original 2-3 tests it remained tentative until a at least 2 out of 3 tests 3 months later continued to show the progression at the same points. By duplicating the testing process the authors were able to minimise the false detection of progression (See section 4.5.3)

1.5.4 Cross-Meridional analysis

All cross-meridional analyses make an attempt to use some of the spatial information contained within a visual field test, and compare points or clusters in the upper hemifield with those in the lower one.

1.5.4.1 Glaucoma Hemifield Test

The Statpac software available for the HFA contains an algorithm called the Glaucoma Hemifield Test (GHT). The GHT automatically evaluates a single threshold test producing a summary using a single line of text (Åsman and Heijl 1992). The visual field is divided up into 10 sectors where test points are arranged along directions taken by nerve fibres. Five sectors are in the superior hemifield, their shape and position mirroring those in the inferior hemifield (figure 13). Using Statpac, the significance of each point's deviation from its expected normal value is calculated. From that significance score a probability score is calculated for each point. The probability scores for points in each sector are summed and the difference between mirror image sectors calculated. In addition a General Height (GH) value is calculated. Points that constitute the 24-2 field are ranked using their age related deviation from normal scores. The 85th percentile (7th highest value) is taken as an index of the general sensitivity of the whole field. The limits of normality for GH and mirror sector differences were calculated by applying the algorithms to a large set of normal visual fields. From the ensuing data a final algorithm is generated that produces a single line of text as a comment on the field. This algorithm will also flag up fields with symmetrical field loss as well as those with diffuse loss of sensitivity. The comment is only intended to be used to discriminate patients with normal or glaucomatous field loss.

Follow up work has looked at the repeatability of the GHT and its usefulness in selecting patients for inclusion into trials of patients with glaucoma or ocular hypertension (Katz et al. 1995). Patients who had been diagnosed as normal, ocular hypertensive or mildly glaucomatous on the basis of examination and static perimetry underwent at least 2 30-2 fields. Fields scored by the GHT to be "outside normal limits", "generalised

reduction in sensitivity" and "borderline/ generalised reduction in sensitivity" were labelled as abnormal. Other GHT scores were labelled as normal. The authors looked at the sensitivity and specificity changes when using the GHT score from one or 2 fields in selecting patients for potential trials that require either ocular hypertensive or glaucomatous patients (table 4). The figures for selecting patients for a glaucoma trial show a small change in sensitivity for a larger gain in specificity if 2 abnormal GHT scores are required.

Table 4 Sensitivity and specificity of single and repeat visual field testing using the GHT to classify visual field test results.

	OH Trial		Glaucoma Trial	
	One Test	Two Tests	One Test	Two Tests
Sensitivity	80.8	69.8	84.2	80.0
Specificity	84.2	89.5	80.8	89.9

Testing ocular hypertensive patients a second time increases the specificity slightly while reducing sensitivity. With glaucoma patients there is a greater rise in specificity with less decrease in sensitivity(Katz, Quigley et al. 1995).

The situation is different in selecting ocular hypertensive patients where there is a larger fall off in sensitivity coupled with a smaller gain in specificity when moving from one to two fields. This may well be due to the ability of automated fields to detect glaucomatous visual loss earlier than manual perimetry (Katz et al. 1995). Subsequent work looking at serial GHT scores in the same group of patients showed that one abnormal GHT score is not a consistent criterion for defining glaucomatous field loss in patients with previously normal visual fields (Katz et al. 1996). Of the 177 ocular hypertensive patients with normal visual fields at entry, 68 subsequently developed an abnormal GHT

score. However only 35 showed a sustained abnormal score over 2 fields; over 3 fields the figure was 25.

1.5.4.2 GLASS mirror image method

Using the assumption that early glaucoma produces asymmetric field loss a simple scoring system has been developed that compares the sensitivities of groups of points above and below the horizontal meridian (Duggan et al. 1985; Sommer et al. 1987). The sum of threshold values in each group is compared with the corresponding group the other side of the horizontal meridian. The criteria for detecting significant differences between groups were arbitrarily varied. Criteria were developed that produced sensitivity and specificity figures greater than 90% in separating early glaucomatous from normal fields. Additional criteria were added to detect diffuse but not symmetrical field loss. This test has not been formally given a name has been referred to as the GLASS mirror image method (Katz et al. 1991) since it is designed to be used in the Glaucoma Screening Study in the USA.

1.5.5 Perimetric nerve fibre bundle maps

Like the GHT, perimetric nerve fibre bundle maps use clusters of visual field test points based on nerve fibre layer defects. However analysis is performed by cluster and not across the horizontal midline. Perimetric nerve fibre clusters may be defined fluidly as they arise in each patient (Åsman and Heijl 1992) or may be a specified arrangement of points derived by matching nerve fibre bundle/ visual field defects to sectors of the optic disc.

Arcuate cluster analysis (Åsman and Heijl 1992) assesses points in relation to the normal nerve fibre layer. If adjacent points are depressed and interconnected by lines corresponding to the normal nerve fibre layer then a cluster is formed. Either the corrected pattern deviation value or a pointwise probability score (Åsman and Heijl 1992) can be used to define whether a point is depressed or not. By summing the corrected

pattern deviation values or pointwise probability scores a total score for each cluster can be calculated. If more than one cluster is present in a field then the one with largest volume is used for analysis.

High resolution perimetry has been used to accurately define scotoma border in eyes with nerve fibre layer defects (Weber and Ulrich 1991). By matching scotoma to nerve fibre layer defect they were able to produce a map of nerve fibre layer equivalent lines in the visual field. From this they suggested a 21 sector division of a 30 degree field, admitting that this would cause some problems given that such an automated field would only test 70 points. In spite of this others have gone on to relate focal visual field defects to focal optic disc changes using the scanning laser ophthalmoscope and either short wave perimetry (Yamagishi et al. 1997) or conventional white on white perimetry (Anton et al. 1998). Recently an alternative cluster map has been derived. It was constructed by superimposing a 24-2 visual field test pattern on nerve fibre layer photographs (Garway-Heath et al. 2000)

1.5.6 Comparison of detection systems

The developers of the glaucoma hemifield test compared it with the GLASS test, assessing the sensitivity and specificity of both tests in discriminating glaucomatous from normal visual fields (Åsman and Heijl 1992). The GHT had a higher sensitivity and specificity than the GLASS test. By modifying the GLASS test it was shown that the differences in sensitivity and specificity between it and the GHT were due to

1) the GHT uses deviations from age related scores or probability scores rather than threshold scores
2) the GHT derives the normal range for differences between sectors purely by using a large database of normal fields rather than training the algorithm on normal and selected glaucomatous fields.

Åsman and Heijl used cluster volume or largest sector up-own difference in logistic regression analysis of normal and glaucomatous fields (Åsman and Heijl 1992). Greater sensitivity and specificity was seen when using pointwise probability scores rather than corrected pattern deviation values to calculate the cluster volumes and sector up-down differences. In using pointwise probability scores account is taken of the empirically measured normal range of cluster volume rather than assuming a fixed range of normality with the pattern deviation scores. Using the measured range of normal variation at each position helped increase discriminative ability of the analysis.

In a large study of early glaucoma and normal patients (Katz et al. 1991) Katz et al measured the sensitivity and specificity of 10 different algorithms in recognising glaucomatous field loss. The algorithms were a mixture of global indices, cluster analysis and cross-meridional analysis including the GLASS method and the GHT. The patients were part of a cohort experienced at kinetic perimetry who then underwent, for most of them, their first automatic field test using the 30-2 program on the HFA. Patients were classified as being normal or having glaucoma on the basis of IOP, manual perimetry and a family history of glaucoma. Table 5 summarises the sensitivity and specificity of the 10 algorithms. Of the global indices only CPSD, P<5% yielded a high sensitivity and reasonable specificity. The sensitivity values in the other 3 global indices were all low. The other 6 cluster and cross meridional techniques all had high sensitivities and reasonable specificity. The GHT and GLASS methods both had higher sensitivity and specificities than in the Åsman study comparing the 2 (Åsman and Heijl 1992) possibly reflecting differences in patient and field selection criteria. The study also found that approximately one third (glaucoma 36.8%, normal 29.3%) of all patients failed to meet the HFA criteria for field reliability; this was mainly due to fixation losses. Some unreliable patients were reclassified as "readable" if either their fixation loss rates or their false negative rates failed to meet reliability while the other 2 reliability rates were within acceptable limits. Highest sensitivity and specificity values were seen in reliable patients with progressively lower

values in the readable and unreliable group. Such a comparative study shows the usefulness of field analysis algorithms in using field data in research studies, especially where hard entry criteria are required. What they cannot do state categorically whether a field is glaucomatous or not. The authors went on to point out that different algorithms may be needed to detect progression compared to those required to detect initial glaucomatous field loss.

Table 5 Sensitivity and specificity of 10 algorithms for detecting glaucomatous field loss.

	Sensitivity	Specificity
Global Indices		
MD, P<5%	77	88
MD, P<1%	68	94
CPSD, P<5%	91	84
CPSD, P<1%	75	92
Clusters		
AGIS	96	78
LTG, dB	92	84
LTG, P values	97	84
Cross-Meridional		
GLASS mirror image	92	88
GHT abnormal only	92	87
GHT abnormal/borderline	93	77

Global indices offer either high sensitivity or specificity. Cluster analysis and cross-Meridional analysis offer better combinations of higher sensitivities and specificities. However for all of the algorithms tested there is a trade off between sensitivity and specificity. None of the algorithms offering >90% sensitivity can produce >88% specificity. This is less important on an individual basis where IOP and disc appearance can be taken into account. However it is important in trials and screening programs where pre-defined outcome measures are used. (Katz et al. 1991).

1.6 Detecting Glaucomatous Visual Field Progression

1.6.1 Observer Opinion

As part of a larger study 6 ophthalmologists were sent at least 4 automated fields from 30 patients (Werner et al. 1988). The fields had been performed over the course of one year. Agreement among the ophthalmologists was poor with only 5 out of the 6 agreeing on whether 15 of the 30 were progressing, improving or showing no change. A later study found poor inter (range of κ= 0.478 to 0.606) and intra-observer (range of κ= 0.583 to 0.715) agreement in assessing field series using 3 ophthalmologists (Nouri-Mahdavi et al. 1997). Despite such problems the Glaucoma Laser Trial used an experienced clinician at a visual reading centre to grade fields as better, worse, unchanged, or better (Glaucoma Laser Trial Research Group 1995). After 2 years of follow-up the rates of progression were 9-13% according to treatment group. Interestingly the rates of improvement were 22%. This may be because no effort was made to minimise learning of visual field technique before the trial started.

1.6.2 Global Indices

Early work followed mean sensitivity (Schultz et al. 1987; Werner et al. 1988) or total threshold (Holmin and Krakau 1980) across all points. Subsequent attempts at statistical analysis of these values compared mean sensitivities using t tests (Werner et al. 1988) or used linear regression of total threshold (Holmin and Krakau 1982; Wu et al. 1986; O' Brien and Schwartz 1990; O' Brien et al. 1991)

The global indices generated by Statpac software for the HFA have been subjected to linear regression (Chauhan et al. 1990; Smith et al. 1996; Katz et al. 1997; Nouri-Mahdavi et al. 1997). There is poor agreement between results from linear regression of global indices and either observer opinion or pointwise linear regression (Chauhan et al. 1990; Nouri-Mahdavi et al. 1997).

Using the CIGTS system (1.5.2.1) an increase in the field score of 3 or more compared to the average score from 2 baseline fields is deemed to constitute progression (Katz 1999). This requires that the increase in score is confirmed by 2 additional tests.

1.6.3 Cluster Analysis

The Advance Glaucoma Intervention Study (AGIS) has recently reported its findings (AGIS 1998). The study randomised participants to either an ALT-trabeculectomy-trabeculectomy (ATT) sequence or trabeculectomy-ALT-trabeculectomy sequence of interventions. Their results have been broken down according to the racial origin of the patients because of "unexpected findings of a statistically significant interaction between race and assigned intervention sequence". Field progression is defined as a change in visual defect score of 4 or more. At 5 years the rates of progression for black patients are 21.3-34.5% (according to treatment intervention) and 13.1-40.6% for white patients. The study authors found that reductions in visual field and visual acuity were greater in blacks assigned to receive TAT than ATT whereas the reverse was true for whites.

Alternative criteria for progression were developed for the Normal-tension Glaucoma Study (Schulzer 1994). The study design required an initial defect to consist of at least 3 contiguous points depressed by 5 dB or more, with at least one point depressed by 10 dB or more. The progression criteria initially proposed required that 2 points within or adjacent to a baseline defect declined by 5 dB or three times the average baseline fluctuation of that eye, whichever was greater. The sensitivity of each point also had to be outside the range of values measured for that point on 3 baseline visits. Any suspected change led to the patient being asked to attend for 2 further fields. If the progression was confirmed in at least one of the fields then a definite diagnosis of progression was made. An early report from the study led to a revision of these criteria (Schulzer 1994). Each individual test was found to have a lower than expected specificity rate of 84.3%. This

meant that even though progression had to be confirmed, the overall false positive rate for detecting it was unacceptably high at 4.5%. It was only by requiring a) the minimum sensitivity change to be 10 dB and b) suspected progression to be confirmed on 2 out of a further 3 fields 3 months later that the false positive rate fell to 2%. In contrast the study had estimated the rate of progression at 0.2% per patients per 3 months.

A comparison of the AGIS and CIGTS scoring systems found that both detect large amounts of progression and improvement (Katz 1999). When comparing the scores from fields obtained 1 year apart, 15% progressed and 20.9% improved using the CIGTS requirement of a change of 3 or more. The figures using the AGIS scoring system were 7.5% showed progression, and 11.9% showed improvement, using a requirement of a change of 4 or more. The rates of change became more comparable when a minimum change score of 3 for AGIS was used.

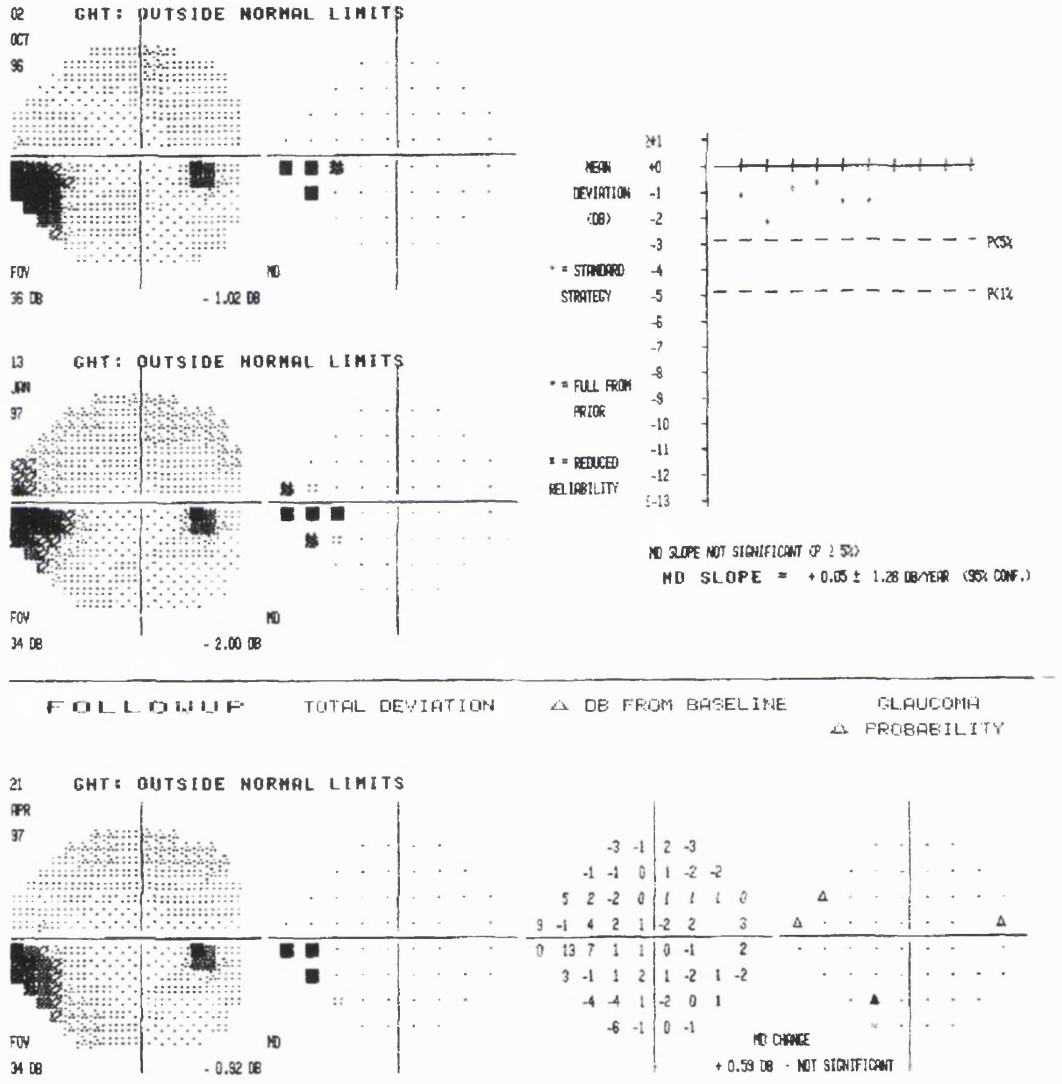
A similar study with on average 6 years follow up found progression rates of 10% and 19% using AGIS and CIGTS respectively. The figures for improvement were 3% and 0% (Katz et al. 1999).

1.6.4 Statpac II Glaucoma Change Probability Analysis

The initial Statpac software was capable of publishing a summary of serial fields. Its analysis was limited to performing linear regression of Mean Deviation (MD) (Heijl et al. 1986). The second version, Statpac II, contains Glaucoma Change Probability software, designed to assess changes in HFA visual fields using normative physiological data (Heijl et al. 1990). Fifty one patients with glaucomatous field loss were tested using the 30-2 threshold strategy 4 times within a 4 week period (Heijl et al. 1989). From this series of fields point-by-point normal values for variability were established. The same study goes on to quantify the pointwise threshold change from one test to the next as a function of deviation from normal at the initial test. They clearly highlight the variability seen in pointwise field analysis. Points initially depressed by 8 to 18 dB on subsequent testing have

a 95% chance of being anywhere from normal threshold to absolute defect. Comparing the mean of 2 initial fields with the mean of 2 subsequent fields reduces fluctuation but does not abolish it, unfortunately no quantification of the reduction in variability is given (Heijl et al. 1989).

Figure 9 Statpac Glaucoma Change Probability Printout



Top left are 2 baseline threshold greyscale plots and their corresponding Total Deviation probability plots. Below them are the corresponding plots for the follow up field. Top right is a plot of Mean Deviation over time with lines indicating the probability that any change was due to chance. The slope of any change in MD is also reported. Bottom right shows the change in sensitivity from the mean of the baseline sensitivities, and the corresponding probability plot. Filled triangles represent sensitivity reduction unlikely to be due to variability ($p<0.05$). Empty triangles represent sensitivity improvement.

When provided with a series of fields the Statpac software generates threshold values for each point based on the mean of the first 2 fields. It will then compare any subsequent field test with the mean field. Points in subsequent test that are outside the normal range of variability are labelled as improving or progressing with a statistical significance of $P<0.05$ (figure 9). The linear regression of MD can be altered to take account of learning. By regressing the field series with and without the first field it is possible to determine whether the first field should be included. If the MD of the first field is significantly out of line with subsequent fields then that field is automatically excluded from all analyses. As in the original Statpac software box-whisker plots showing the distribution of pointwise deviation from expected normal thresholds are also available. The benefit of Statpac 2 is that it detects change above that seen in a series of non-progressing glaucomatous fields. Its disadvantages are:

It is reliant on the original database of visual fields for calculation of normal variability. For example confidence intervals for progression are only valid for those points with an initial deviation of –12 dB to +2 dB.

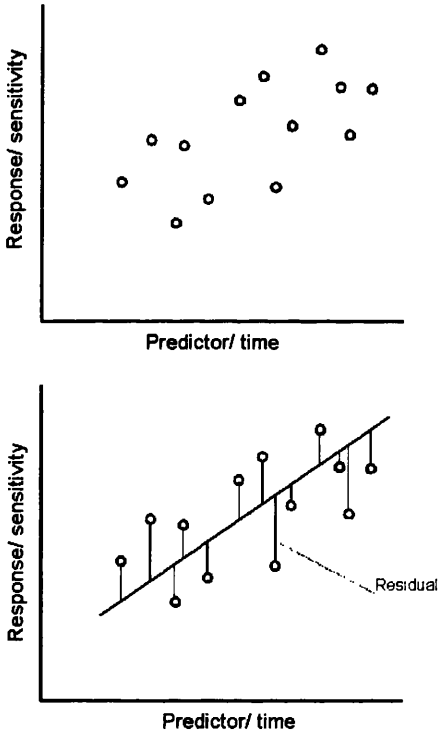
It does not use all the information in a field series; instead it compares the final field (although another can be specified) with the mean of the first 2. Any changes seen in the intervening fields are ignored.

1.6.5 Linear regression

The use of linear regression to analyse visual field data has proved attractive because it uses all the information from each field test in a series. The analysis describes the relationship between continuous variables. One variable is deemed to be the dependant or response variable (sensitivity to light in most situations in this thesis), while the other is the predictor or independent variable (time in this thesis). The analysis attempts to fit a straight line to the data that gives the best prediction of the response variable from the predictor variable (figure 10). The "fit" of the line is judged by

calculating the sum of the squares of the vertical distance of each point from the proposed line. The vertical distance of each point from the line is known as the residual and may be positive or negative depending on whether it is above or below the line. Squaring the residuals avoids sign problems and deals purely with the distance of the points from the line and not whether they are above it or below it. The line produced by the analysis is the one that minimizes the sum of the squares of the residuals.

Figure 10 Regression line fitted to data set so as to minimize the squares of the residuals



A proposed line is drawn through the points under study. The vertical distance from each point to the line is known as the residual. The equation defining the exact position of the line is calculated in such a way so as to minimise the sum of the square of the residuals.

The calculations in linear regression involve some assumptions about the data (Altman 1991):

1) The values of the outcome variable should have a Normal distribution for each value of the predictor variable.
2) The variance or standard deviation of the outcome variable should be the same for each value of the predictor variable.
3) The relation between the 2 variables under test should be linear.

The estimated slope b is given by the equation

$$b = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2}$$

$\bar{x}$ and $\bar{y}$ are the means of the predictor and outcome variables respectively.
x_I and y_I are the observed data and where (i =1,2,3,.....n)

The calculations involved in linear regression can be simplified in their expression by calculating the sum of the squares of the predictor and outcome variables about their means and the sum of products.

$$S_{xx} = \sum x_i^2 - \frac{(\sum x_i)^2}{n}$$

$$S_{yy} = \sum y_i^2 - \frac{(\sum y_i)^2}{n}$$

$$S_{xy} = \sum x_i y_i - \frac{\sum x_i \sum y_i}{n}$$

slope can now be expressed as

$$b = \frac{S_{xy}}{S_{xx}}$$

The significance level or "P value" of the slope is derived from a hypothesis test that b=0. The null hypothesis is that there is no relation between the predictor and outcome variables. Dividing the slope by the standard error of the slope and comparing the result with the t distribution with n-2 degrees of freedom tests this hypothesis.

$$t = \frac{b}{se(b)}$$

$$se(b) = \frac{S_{res}}{\sqrt{S_{xx}}}$$

where S_{res} is the square root of the sum of the square of the residuals.

$$S_{res}^2 = \frac{1}{n-2}(S_{yy} - bS_{xy})$$

1.6.5.1 Use of linear regression in field progression

Since the introduction of linear regression to analyse automated visual field data in the 1980's work has looked at point, sector, and fieldwise linear regression changes within a field series (Wu et al. 1986; Chauhan et al. 1990; O' Brien and Schwartz 1990; Noureddin et al. 1991; O' Brien et al. 1991; Smith et al. 1996; Bhandari et al. 1997; Katz et al. 1997), or has compared PLR to Statpac 2 and other analytical techniques in the detection of change (Birch et al. 1995; Nouri-Mahdavi et al. 1997; Viswanathan et al. 1997).

Schwartz's group used what they termed "trend analysis" to follow a threshold variable over time (Wu et al. 1986; O' Brien and Schwartz 1990; O' Brien et al. 1991). The threshold variable used was the mean value for the whole field or sector of the field, or the threshold of each point. Their use of this analysis used linear regression to calculate the slope of threshold over time, and the Spearman P value of the correlation coefficient of threshold values over time to calculate whether there is zero correlation. Although this sounds dissimilar to pure linear regression the mathematics testing the null hypothesis of zero correlation are exactly equivalent to that for the hypothesis of zero slope in linear regression analysis (Altman 1991). The same group presented results from whole field and sector analyses rather than from pointwise analysis. They described regional variation in rates of field loss with greatest rate of loss in the temporal and superotemporal regions (O' Brien and Schwartz 1990). Furthermore they described a weak association between IOP and rate of change in 1 of 7 possible sectors of the visual field (O' Brien et al. 1991). Subsequent use of pointwise linear regression to analyse the fields of patients with normal tension glaucoma (NTG) found no difference in IOP between those patients deemed progressing or stable. However a study looking at NTG patients undergoing trabeculectomy showed that the mean slope of progressing points was reduced when the intraocular pressure was surgically reduced (Bhandari et al. 1997).

More recently software has been developed to graphically display the change in sensitivity at each retinal location tested on a Humphrey visual field (24-2 or 30-2) (Fitzke

and McNaught 1994). The "Progressor" software displays a small bar chart at each point tested. The bars of the chart represent the sensitivity at that point on each test occasion. Colours are applied to the bars to indicate the statistical significance of any change in retinal sensitivity, if any. STATPAC 2 and Progressor have been compared in the time required to detect progression. A series of visual fields from 19 patients with normal tension glaucoma deemed by observers to be deteriorating were analysed. Both programmes were able to detect progression. However Progressor consistently detected change earlier than STATPAC 2 (Viswanathan et al. 1997). It should be noted that the criteria for change were 1) Progressor: 1 point changing with a slope of >1 dB/yr, $P<0.05$ 2) Statpac: 1 point showing significant deterioration from baseline on 3 consecutive occasions.

Other comparative analyses have looked at fields that have not had any observer opinion passed over them. Pointwise linear regression has been compared with other criteria in the frequency of detection of glaucomatous field progression. Three large studies compared the ability of PLR to detect change against 1) observer opinion, linear regression of MD and Glaucoma Change Probability (both from Statpac II) (Birch et al. 1995) 2) linear regression of MD, CPSD, and mean threshold of clusters corresponding to the glaucoma hemifield test (Katz et al. 1997), and 3) linear regression of Statpac indices, mean threshold of clusters, observer opinion, and multivariate analysis (Nouri-Mahdavi et al. 1997). What is clear is that although there is some degree of agreement by different techniques there are always a considerable proportion of patients detected as progressing by one technique and not by the others. What is also critical in examination of detection of progression are the criteria used to define change and the patients tested. Most studies using linear regression have concentrated on patients with glaucomatous field loss (Birch et al. 1995; Smith et al. 1996; Bhandari et al. 1997; Katz et al. 1997), however there are some that also looked at ocular hypertensives as well as glaucoma patients (Birch et al. 1995). Table 6 summarises some of the criteria used in pointwise linear regression analyses.

Some studies using pointwise linear regression analysis have used a single changing point as the criterion for change (Noureddin et al. 1991; Birch et al. 1995; Smith et al. 1996; Katz et al. 1997). Others have scored the mean incidence per patient of points changing with under varying criteria (Wild et al. 1997). Others have used a single changing point as the entry criterion (Bhandari et al. 1997). Although a single point has often been used to indicate change, the criteria for change at that point have varied with different slopes, levels of significance, presence of progression in successive fields. Because linear regression analysis is performed at up to 74 locations in a single field the risk of chance progression becomes real. Accordingly slopes different from zero have only been considered by some to be significant if the P value is less than 0.001. This changes the overall P value to a value closer to 0.05 (Smith et al. 1996; Katz et al. 1997). Others have used 0.01, and 0.05 (Birch et al. 1995). A variety of critical slopes have been used as part of the definition of a changing point:

- any non zero significant slope (Smith et al. 1996; Katz et al. 1997)
- 0.2 dB/month (Noureddin et al. 1991),
- 1dB/year for the whole field (McNaught et al. 1996; Bhandari et al. 1997; Viswanathan et al. 1997),
- a slope greater than the normal age-decline for each point (Wild et al. 1997).

Where a figure of 1 dB/yr has been used, it has been selected because a) it represents approximately 10 X the normal age related decline (

figure 6), and b) a point with a normal sensitivity of 30 dB would take 30 years to loose all sensitivity.

A large retrospective study of visual fields culled from a database (Smith et al. 1996) examined patients who had had at least 7 fields over at least 4.5 years. Univariate linear regression was performed on mean deviation, CPSD, mean thresholds of clusters corresponding to the Glaucoma Hemifield Test (GHT), and thresholds of 52 individual test locations. Less than a third of all subjects showed deterioration using any technique

over a mean follow-up of 7.1 years. Where a technique detected progression the rate of change is 1-5 dB /year depending on the technique used. For series where the slope of MD change was significantly less than zero the mean slope of MD change was –1.26 dB/yr. The figures for the clusters that compose the glaucoma hemifield test ranged from –1.51 to –2.84 dB/yr depending on the cluster chosen. No relation between cluster location and likelihood of progression was seen. The mean slope of progressing points ranged from –0.84 to –5.84 dB/yr, this is equivalent to 14 to 73 times the amount expected due to age. It was felt that a minimum of approximately 5 years of follow-up with annual perimetry were required to detect significant changes in the fields by linear regression.

The limited number of field tests that patients undergo means that a large amount of noise is detected. Alternative types of regression analysis such as "pointwise multivariate regression analysis with fixed effects on panel data" (RAFEP) have been proposed in an attempt to deal with this situation (Nouri-Mahdavi et al. 1997). However the detection rates for RAFEP were no different from those obtained using pointwise linear regression. In a study using predetermined criteria for visual field progression regression, analysis of global indices and indices derived from clusters of depressed points failed to accurately separate progressing and non-progressing fields (Chauhan et al. 1990). Wild has modelled the visual field topographically and longitudinally. The longitudinal component was determined using pointwise linear regression (Wild et al. 1993).

Linear regression of the mean sensitivity of points in clusters has been performed for Octopus (Wu et al. 1986; Werner et al. 1988; O' Brien and Schwartz 1990; O' Brien et al. 1991) and the HFA (Smith et al. 1996; Katz et al. 1997). None of these studies required a minimum negative slope to define progression. The same group performed both studies on the HFA and for both the predetermined P value for significant change in each of the clusters was $P<0.005$. This Bonferroni correction made the overall P value for detecting a non-zero slope equal to 0.05. The correctionadjusts the P value to allow for multiple tests of significance (Altman 1991). This correction was not made in the Octopus studies.

Table 6 Change criteria used in studies of pointwise linear regression

Reference	Number of fields	Years of follow up	Progression Criteria			
			Number of changing points	Minimum slope (dB/yr)	P Value	Comment
Noureddin (Noureddin, Poinoosawmy et al. 1991)	8	2.3	1	-2.4	<0.05	
Birch(Birch, Wishart et al. 1995)	6	3.3	2	-1	Not known	
Smith(Smith, Katz et al. 1996)	9.5	7.1	1	<0	<0.001	
Katz(Katz, Gilbert et al. 1997)	6	5	1	<0	<0.001	
Nouri-Mahdavi(Nouri-Mahdavi, Brigatti et al. 1995)	6.8	5.6	2 contiguous	<0	<0.0009	
Viswanathan (Viswanathan, Fitzke et al. 1997)	Not known	Not known	1	-1	<0.05	
Bhandari (Bhandari, Crabb et al. 1997)	9.7,8.6	Not known	1	-1	<0.001	Pre and post op fields analysed
Wild(Wild, Hutchings et al. 1997)	12	6	1	<Normal age decline	Not known	Also 2.5,5, 10 x normal age decline tested

Previous studies using pointlwise linear regression criteria have used progression criteria that have required either 1 or 2 points to be changing, with a minimum slope ranging from 0 to -2.4 dB/yr, and with a P value that was either not specified or when specified ranges from P<0.05 to <0.0009.

1.6.5.2 Spatial relationships and PLR

Points within a visual field are not independent, some correlation exists between point thresholds up to a separation of 30° (Lachenmayr et al. 1995). However the relationship between points remains poorly understood. It has been claimed that variability increases with eccentricity (Heijl et al. 1989). However having corrected for the effect of eccentricity on initial sensitivity others have found that variability is not eccentricity dependant (Boeglin et al. 1992). Furthermore the correlation of threshold variation decreases as the distance between any 2 points increases (Heijl et al. 1989).

Equations have been generated that describe sensitivity at a point in the visual field given a general stimulus location (x,y) (Wild et al. 1993). Mathematical models that take in to account the spatial dependence that exists between points in a glaucomatous visual field have been developed (Crabb et al. 1996).

Linear regression in its simplest form performs analyses as though all points are unrelated. From our knowledge of the topographical arrangement of the retina and from clinical experience it would seem unlikely that glaucomatous damage would occur at random unrelated points within a visual field. In an analysis of progressing points using linear regression it has been shown that in field series with ≥2 progressing locations 86% of the locations are adjacent (Smith et al. 1996). Work that has used contiguous points as part of progression criteria represents an attempt to look at spatial relationships in field data (Nouri-Mahdavi et al. 1997; Wild et al. 1997).

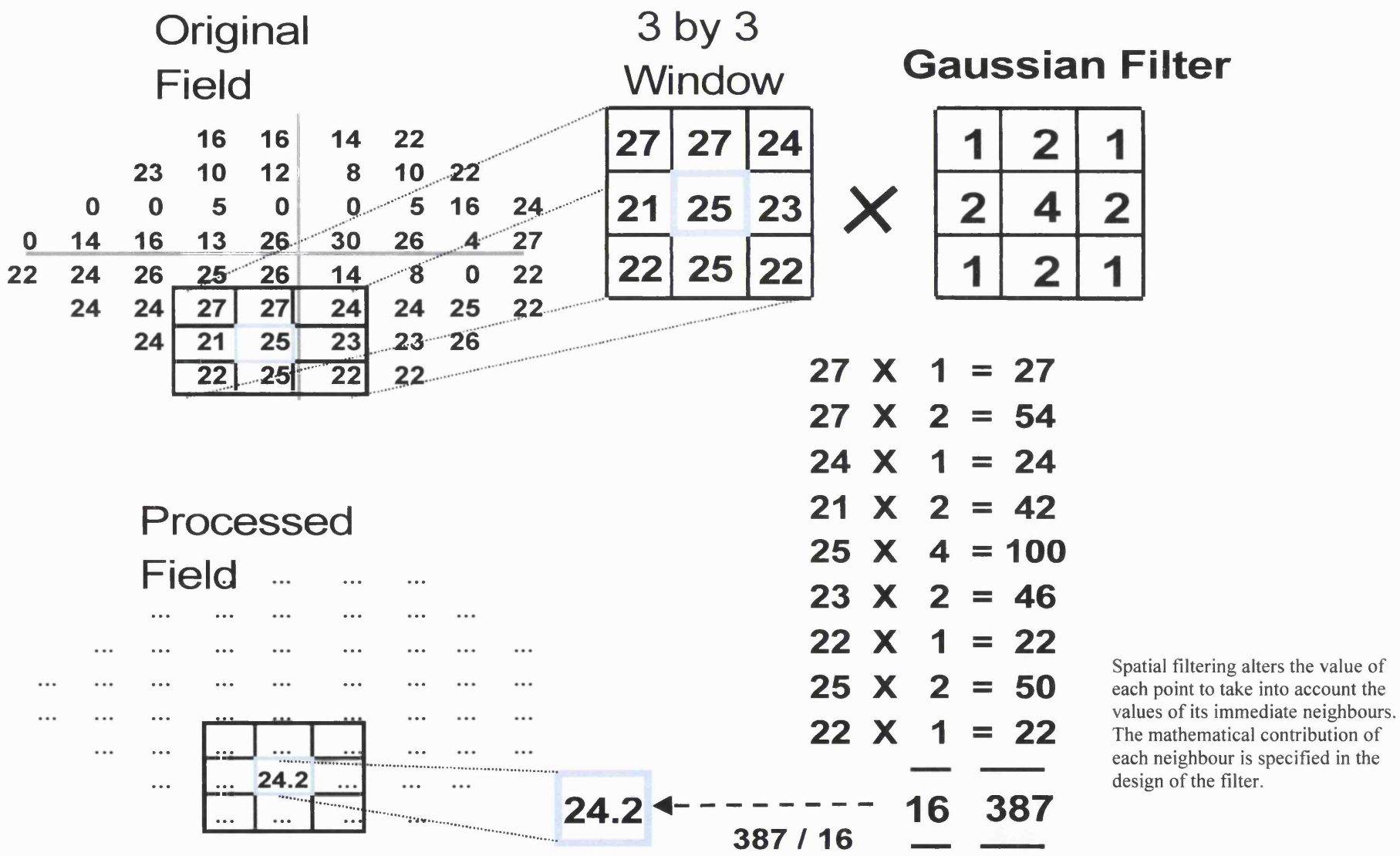
The individual point threshold values from an automated visual field are arranged in a grid. Such a grid of numbers is no different to that of a digital image. Image processing has been used to improve the quality of digital images and similar techniques are now being applied to visual field images. Several papers have used spatial filters to try and remove noise (Crabb et al. 1995; Fitzke et al. 1995; Viswanathan et al. 1996; Crabb et al. 1997). Furthermore by comparing filtered and unfiltered fields a new estimate of variability in data has been derived (Crabb et al. 1995). Indeed a simple filter is available as part of the

Progressor software. Not only does this technique reduce noise but also relates the values at each point in a field to those of the points surrounding it. The technique involves moving a 3x3 window over the grid of visual field threshold values. The centre of the window passes over each point. Each threshold value falling within the window is multiplied by the weighting value in that cell of the window. These values are then averaged to generate a new, filtered value (Figure 11). If one assumes that visual field progression is linear then one can project the slope values from linear regression forward to predict future pointwise field values. Using such a spatial filter improves the prediction of visual field progression using linear regression (Crabb et al. 1997). Crabb's study analysed 5-field series and then predicted the pointwise values at 1 and 2 years after the initial fields. A filter improved the prediction precision from 56% to 73% at the 2 year mark. However what has not been shown is that using spatial filtering detects change in glaucomatous visual field earlier (Viswanathan et al. 1996). Hopefully knowledge of the spatial dependence of pointwise sensitivity across the glaucomatous visual field will enable customised spatial filters for each point in the visual field to be derived (Crabb et al. 1996).

1.6.6 Neural Networks

Very little work has been done using neural networks on visual fields. Their inherent attraction is that they could detect glaucomatous visual field changes without necessarily having to be fed rules about what constitutes such change. One study has shown that a neural network can organise field defects into different types (Henson et al. 1996) while another has shown how networks can learn to separate fields with early glaucomatous damage from normals (Brigatti et al. 1996). However no study has looked at the ability of neural networks to detect visual field progression. Furthermore any studies using this technique will be limited by the fact that there is no gold standard to compare against. Since neural networks learn from the material fed to it, it seems unlikely that a simple network will detect glaucomatous damage earlier and more reliably than humans.

Figure 11 Spatial filtering of a visual field



Spatial filtering alters the value of each point to take into account the values of its immediate neighbours. The mathematical contribution of each neighbour is specified in the design of the filter.

1.6.7 Field conversion in OHT

An early study using automated perimetry used a mixture of 4 types of perimeter to follow ocular hypertensives that were randomised to either treatment or no treatment (Schulzer et al. 1991). Criteria for an abnormal field were the development of any one of the following:

a) 2 adjacent points depressed by 6 dB or more

b) A single point in the central 30 degrees depressed by 10 dB or more

c) A point within 5 degrees of fixation depressed by 0.4 dB or more

d) Adjacent peripheral points depressed by 10 dB or more provided that it wasn't part of a generalised depression

Any abnormality confirmed on the Perimetron or Peritest had to be confirmed on the an Octopus 201 or HFA. No justification for these criteria was given.

Early studies using automated perimetry to assess conversion in ocular hypertensives were complicated by the fact that the perimeters were introduced after the study began. In a treatment/ no treatment study of ocular hypertensives (Kass et al. 1989) an abnormal automated field was defined as having 4 contiguous points depressed by 5 dB or more. No mention was made of whether points may be contiguous across the horizontal midline. In addition defects had to be reproducibly detected in the same area on 3 successive visits. A trial with similar inclusion criteria also introduced automated perimetry after the onset of the trial (Epstein et al. 1989). The authors found that many of their patients who were classifiable as normal on Goldmann perimetry had abnormal Octopus fields. Unfortunately no criterion for abnormality was given.

1.6.8 Non-Glaucomatous Visual Field Change

In the last decade the automated perimeter has started to make inroads in to the neuro ophthalmology clinic. Attempts to monitor patients using such perimetry have relied on using simple comparisons of fields. A study evaluating the effects of optic nerve sheath

fenestration merely compared the total sensitivity loss before and after surgery (Pearson et al. 1991). In the Optic Neuritis Treatment Trial (ONTT) a series of post intervention fields were compared with pre-intervention fields using i) Statpac indices and ii) a grading system that divided fields as showing diffuse or localised loss according to whether the majority of a defect on the overall probability deviation plot persisted on the pattern deviation plot (Keltner et al. 1993; Keltner et al. 1994). Recent work has pointed out the large amount of long and short term variability in patients with optic neuritis compared to normals. The authors make the point to distinguish genuine sensitivity changes from variability requires "more than a comparison of the current visual field with the most recent previous visual field" (Wall et al. 1998). This may explain why the definition in the ONTT of a visual field returning to "normal" was when it had normal global indices and less than 8 of the 76 individual thresholds were outside normal limits (Keltner et al. 1994).

1.7 Interventions in Glaucoma

Despite the caveats of section 1.2.1 most treatment options in glaucoma have been developed as a means of reducing intraocular pressure. This is because it is the most easily modifiable risk factor. Other major risk factors such as age, gender, race and family history are not modifiable (Leske et al. 1995). The choices facing an ophthalmologist are:

a) Drops

   Beta-blockers

   Carbonic anhydrase inhibitors

   Muscarinic agonists

   Alpha agonists

   Prostaglandin analogues

b) Systemic suppressers of aqueous production

   Carbonic anhydrase inhibitors

c) Argon laser trabeculoplasty

d) Filtration surgery

Trabeculectomy

Drainage tube procedures

e) Ciliary body ablation

Cyclocryotherapy

Cyclophotoablation

Initially it was thought that any glaucoma treatment would act solely by reducing aqueous production or increasing aqueous outflow and thus reducing intraocular pressure. However many patients with glaucoma do not have raised IOP (Sommer et al. 1991) or continue to develop glaucomatous damage after their IOP is reduced to apparently safe levels (Brubaker 1996). This has led to the idea that additional risk factors are at work. Alternative pressure independent systemic stimuli may be responsible for triggering neuronal death (Drance 1972) or factors released by damaged neurons may act locally to perpetuate the degeneration (Yoles and Schwartz 1998).

Systemic vasospastic disorders such as Raynaud's phenomenon and migraine have been associated with normal tension glaucoma (Corbett and Phelps 1985; Phelps and Corbett 1985). This has led to theories that vasospastic mechanisms trigger neuronal death. Indirect retrospective evidence has suggested that patients receiving calcium channel antagonists for non ophthalmic conditions may experience less glaucomatous progression (Netland et al. 1993). This has not been supported by other workers (Liu et al. 1996). An interim analysis of a prospective study has suggested that Brovincamine a calcium channel antagonist can lead to improvement in the visual fields of patients with normal tension glaucoma (Sawada et al. 1996). A small study found no effect from using Nimodipine on standard automated visual field indices in NTG patients (Piltz et al. 1998). However the authors did claim that there was an effect on colour vision.

Other work has suggested topical betaxolol and brimonidine have some neuroprotective effects in addition to their pressure lowering effects. The drugs are known to be capable of reaching the optic nerve in therapeutic quantities when applied topically (Burke and Schwartz 1996; Osborne et al. 1997). Two separate small studies have shown that topical timoptol reduces pulsatile ocular blood flow while betaxolol does not (Boles Carenini et al. 1994), and that patients treated with betaxolol showed greater improvements in their visual fields than those on timoptol (Kaiser et al. 1994).

No work has shown that trabeculectomy or antimetabolites exert an effect (if any) by any means other than reducing intraocular pressure. Many studies have shown a beneficial effect in using antimetabolites either during (Egbert et al. 1993; Robin et al. 1997) or following trabeculectomy (FFSG 1996).

1.8 Justification and Aims

The discussion of the above research shows how the concepts for defining glaucomatous change in visual fields have been developed. Many of the techniques have inherent merit to them. The wide variety of techniques used highlights the uncertainty remaining in our ability to label a field as changing. The lack of a gold standard for visual field progression makes the problem harder to solve. A particular problem is the large amount of threshold fluctuation seen in visual field testing. As Katz has shown, using established change criteria; improvement can be detected in glaucoma trials over long periods of time (Katz et al. 1999). Pointwise linear regression (PLR) is a technique that seeks to deal with the noise present in the threshold measurements at each point in a visual field. Even so the change criteria used when the technique has been applied have largely been empirically chosen. Indeed the visual field progression criteria for the MRC 5-FU filtration surgery study are based on empirical criteria. The aims of this thesis therefore are to:

1) Use existing PLR criteria to describe the amount of change detected in the visual fields of 56 eyes prospectively followed in a glaucoma trabeculectomy trial.
2) Develop new PLR change criteria, and apply them to the same data.
3) Assess whether ocular changes brought about by trabeculectomy could be correlated with visual field changes.
4) Apply the field progression criteria developed from the initial 56 eyes to additional data derived from extended follow up of the original eyes and from subsequent eyes entering the trial.

2 Method

Clinical data for this thesis was obtained from the Medical Research Council (UK) 5-FU filtration Surgery Study based at the Institute of Ophthalmology and Moorfields eye Hospital, London.

2.1 MRC 5-FU Treatment Trial

The Medical Research Council 5-Fluorouracil Trabeculectomy Trial is a prospective, randomised, double masked study on the effect of a single intraoperative application of 5-FU versus placebo in patients at low risk of failure undergoing trabeculectomy. The primary outcome measures are:

- Visual function as measured with Humphrey visual field testing. Analysis of fields using pointwise linear regression analysis.
- Optic disc cupping measured with 3-dimensional analysis using the scanning laser ophthalmoscope and simultaneous stereo photography.
- Intraocular pressure defined as the percentage of patients in each treatment arm who have pressures ≤ 21 mmHg, ≤ 19mmHg and ≤ 14mmHg off medication.
- Incidence of short and long term surgical complications.

The trial is still running at Moorfields Eye Hospital, London. Only patients from Moorfields and its outreach clinics are recruited in to the study. Patients listed to undergo trabeculectomy for the first time are eligible for inclusion in the study if they satisfy the following inclusion and exclusion criteria:

2.1.1 Inclusion Criteria

Patients who are deemed to have inadequate intraocular pressure control with or without medical treatment, who are listed for trabeculectomy by their consultant. They must satisfy the following criteria:

An intraocular pressure ≥ 22 mmHg on at least one visit before the time of listing for surgery and while on the current drop regime.

A Humphrey 24-2 visual field with 2 locations > 5dB less than age-corrected normal or one location >10 dB less than age-corrected normal on the numeric pattern deviation plot. The subject must be able to perform the field test with <20% false positives, <33% false negatives and <20% fixation losses. We felt that with supervision and the ability to repeat tests that we could use stricter reliability indices than those suggested by the manufacturer (1.3.3).

An optic disc which in the opinion of the patient's consultant shows glaucomatous changes as shown by focal or diffuse optic rim thinning.

2.1.2 Exclusion Criteria

1. Other eye already in the study
2. Any previous intraocular surgery
3. Previous glaucoma filtering surgery
4. Aphakia
5. Uveitis
6. Anterior segment neovascularisation
7. Retinal or optic nerve neovascularisation
8. Any intraocular epithelial ingrowth
9. Inability or unwillingness to give informed consent
10. Inability or unwillingness to return for postoperative follow up as prescribed in the trial regimen
11. Unwillingness to accept randomisation
12. Patient less than 40 years of age- thought to be at higher risk of failure
13. Any previous anticancer treatment

14. Any other disease causing visual field loss or likely to cause field loss over the next three years e.g. Diabetic retinopathy, pituitary disease or stroke
15. Pregnancy or female of childbearing age who may be pregnant at time of treatment. A pregnancy test is performed on all women of childbearing age to rule out pregnancy.
16. Cataract that is deemed significant enough to require surgery during the course of the trial or that makes field testing or optic disc recording by either photography or the scanning laser ophthalmoscope unreliable or not technically possible.
17. Patients receiving systemic anticoagulant treatment
18. Any medical condition likely to prevent the patient from regularly attending for the next three years
19. Previous conjunctival surgery at proposed site of trabeculectomy
20. Chronic use of topical or systemic steroids
21. Previous squint surgery

From September 1995 to February 1998 I was employed as the Research Fellow on the MRC 5-Fluorouracil Filtration Surgery Study based at the Institute of Ophthalmology and Moorfields Eye Hospital, London, UK. I worked as part of a team (2 optometrists and a secretary) recruiting patients onto the study and following them up. All patients were seen by myself on each visit. Patients were normally approached at the time of listing in clinic. A small proportion of patients were identified by examining consultants surgical diaries. The trial was explained to the patients who were then given an information data sheet to read. They were telephoned a week later for a decision on whether they wanted to take part in the trial. Those that agreed to take part came to the hospital at least once before their operation in order for baseline tests to be performed. The schedule of visits and tests, relevant to this thesis, conducted during the trial is shown below (table 7). If additional field tests were required preoperatively then the patients either made a second preoperative visit or the tests were performed on the morning of surgery. The patients were examined at least on day 1, week 1, week 2, week 4, week 6, month 3 and at 3

monthly intervals until the end of the 1st post operative year. Visits were 4 monthly for the second post operative year. Extra visits were arranged if clinically indicated.

Table 7 Tests, relevant to this thesis, performed on each visit of the MRC 5-FU trial.

Test	Refraction	Visual Acuity	IOP	HFA 24-2	Modified LOCS Grading
Preop Visit 1	√	√	√	√	√
Preop Visit 2			√	√	
Surgery					
Day 1		√	√		
Week 1		√	√		
Week 2	√	√	√		
Week 4		√	√		
Week 6	√	√	√		
Week 12	√	√	√	√	
Month 6	√	√	√	√	
Month 9	√	√	√	√	
Month 12	√	√	√	√	√
Month 16	√	√	√	√	
Month 20	√	√	√	√	
Month 24	√	√	√	√	√

The trabeculectomy was performed by the patients' consultant or a junior at the grade of registrar or higher. A standardised trabeculectomy technique, based on Cairn's original description, was performed in all cases (Cairns 1969). The operation was sited at 12 o'clock and used a limbus based conjunctival flap. After dissecting the conjunctival flap 5-FU (50 mg/ml) or placebo was applied on pre cut 7x4 mm sponges placed between sclera and conjunctiva for 5 minutes. The treated area was copiously irrigated before a scleral flap 4mm ± 2mm X 4mm ± 2mm was measured out and cut. The operation then proceeded as per Cairns' original description. A member of the trial team was present at the operation to deal with randomisation and to collect clinical data. The randomisation of patients to either intraoperative 5-FU or placebo application did not affect the collection of any clinical data. Analysis of data based on intraoperative treatment will be the subject of

future analyses. The trial protocol restricts all such analyses until at least one year after all trial patients have been recruited and operated on.

Clinical trial data was entered at collection into a booklet. At the end of each visit the relevant page was removed form the booklet and the data entered into a purpose developed database (Access 97, Microsoft Corp, Seattle, USA).

2.2 Clinical Examination/Tests

2.2.1 Refraction

Refractions were performed to determine patients' distance refractive errors and their best corrected distance and near visual acuities. A refraction was performed before each visual field test except for the second preoperative field test where the refraction from the first field was used. The subjective refraction obtained at the previous visit is available. Information on prior visual acuity or treatment was not known. Two research optometrists, under standardised reproducible conditions performed the refractions.

If a patient wore contact lenses then they were removed at least 20 minutes before testing. Corrected distance monocular visual acuities were tested on the Bailey-Lovie retro-illuminated distance chart with a chart-patient distance of 3.2 metres. This test distance was corrected by adding a spherical addition of - 0.30 dioptres to the final subjective result.

The refraction was started using retinoscopy to obtain an objective value of the patient's refraction and as a guide to performing a subjective refraction.
A subjective refraction, in negative cylinder format, was then performed using well fitted trial frames and reduced aperture trial lenses, the eye not under test was occluded. The subjective refraction was conducted with room lights on.

The near addition was determined by adding the minimum positive sphere to the distance correction to obtain the best near visual acuity on the near Bailie-Lovie chart at a working distance of 31.5cm. Back vertex distances was measured for prescriptions >+/- 5D.

The following were recorded at the end of the refraction for the trial eye:

a) Distance prescription

b) Best corrected near and distance logMAR visual acuity

c) Back vertex distance if obtained

2.2.2 Visual Acuity

Corrected near and distance visual acuity were measured using the Bailey-Lovie logMAR chart (Bailey and Lovie 1976). The logMAR visual acuity testing protocol for this study is adapted from protocols used in other studies including:

- The Early Treatment of Diabetic Retinopathy Study (ETDRS) (ETDRS 1985; ETDRS 1991)
- Diabetics Control and Complications Trial (DCCT) (DCCT 1998)
- Macular Photocoagulation Study (MPS) (MPS 1982)
- Prospective Evaluation of Radial Keratotomy (PERK) (Waring et al. 1983)
- Longitudinal Optic Neuritis Study (LONS), in a form similar to the NIH ocular hypertension study (ONTT 1991)

The logMAR visual acuity scale was used because it facilities statistical analysis, while still permitting simple conversion to Snellen nomenclature.

2.2.2.1 Distance visual acuity assessment

Where the visual acuity was tested as part of the refraction procedure then the best correction from that procedure was used. The test was performed with a patient-chart distance of 3.2 metres, with the clinic lights on. The examiner made sure that the patient did not lean forward and thus minimise his/her distance from the chart. The patient was asked to slowly read each letter at a rate of approximately one per second. When the patient could not read a letter he/she should was encouraged to guess. If the patient identified a letter as one of two or more letters, he/she was asked to choose one letter.

The test stopped when it became evident that no further meaningful readings could be made.

Once the patient had identified a letter, it was scored as either right or wrong. Corrections of mis- identified letters could only be accepted if the patient had corrected themselves before moving on to the next letter. The score for each visual acuity assessment was calculated by adding the scores for each line completely read, to the scores for the individual letters read on incompletely identified lines.

2.2.2.2 Near Visual acuity assessment

Near visual acuity was measured in the post-operative period from 2 weeks onwards. The procedure for distance visual acuity was repeated for near using the near Bailie-Lovie chart at a distance of 31.5 cm, with a near correction obtained from the most recent refraction.

2.2.3 Intraocular Pressure

Intraocular pressure was measured using Goldmann applanation tonometry. If the corneal astigmatism was greater than 3.0 D, the prism was rotated so that the axis of the minus cylinder on the prism graduation corresponded to the red mark on the prism holder. Slit lamp magnification was set at 10X. The light source was positioned at an angle of approximately 45° and the aperture maximally opened. A Cobalt blue filter is employed. One drop of premixed fluorescein and benoxinate 0.5% anaesthetic is instilled. If it was necessary to hold the eyelids open, the eyelids were held open against the orbital rim avoiding pressure on the globe. The tonometer measuring drum was set at 10 mmHg. The tip of the prism was then gently brought into contact with the centre of the cornea. While viewing through the slit lamp eyepieces the mires are well focused, centred horizontally, and positioned vertically so that they had equal circumference above and below the horizontal dividing line. The measuring drum was adjusted until the inner borders of the two mires just touch each other or, if pulsation was present, until the mires separated a

given distance during systole and overlapped the same distance during diastole. The prism was then removed from the cornea and the procedure repeated until two successive measurements within 1mm Hg were obtained. Before each measurement the drum was reset to 10 mmHg. The last 2 measurements were recorded. The calibration of the Goldmann applanation tonometer was checked every three months. Analysis of IOP at baseline, and at 12 and 16 months using the Kolmogorov-Smirnov test showed that the values were normally distributed. Baseline IOP values were thus compared with those at 12 and 16 months using a paired t test.

Mean IOP's over 12 and 16 months were also calculated for each patient. When performing the calculation it was assumed that there was a linear change in IOP between each measurement.

2.2.4 Grading of Lens Opacity

Lens opacity was graded preoperatively and annually thereafter. After pupillary dilation, the opacities, of the anterior subcapsular, anterior cortical nuclear, posterior cortical and posterior subcapsular regions of the lens were graded using a modified version of the LOCS III grading system (Chylack et al. 1993). One of the major differences between LOCS II (Chylack et al. 1989) and LOCS III is the change from integer to decimal scaling. Because of a misunderstanding we graded lens opacities by matching them to the nearest integer reference picture on the LOCS III chart (figure 12). This meant that the mildest opacity was graded as I instead of 0, there being no grade 0 reference picture on the chart. All opacity grades were therefore expressed to the nearest integer, no attempt was made to grade them to the nearest 0.1 as was intended by the developers of the chart.

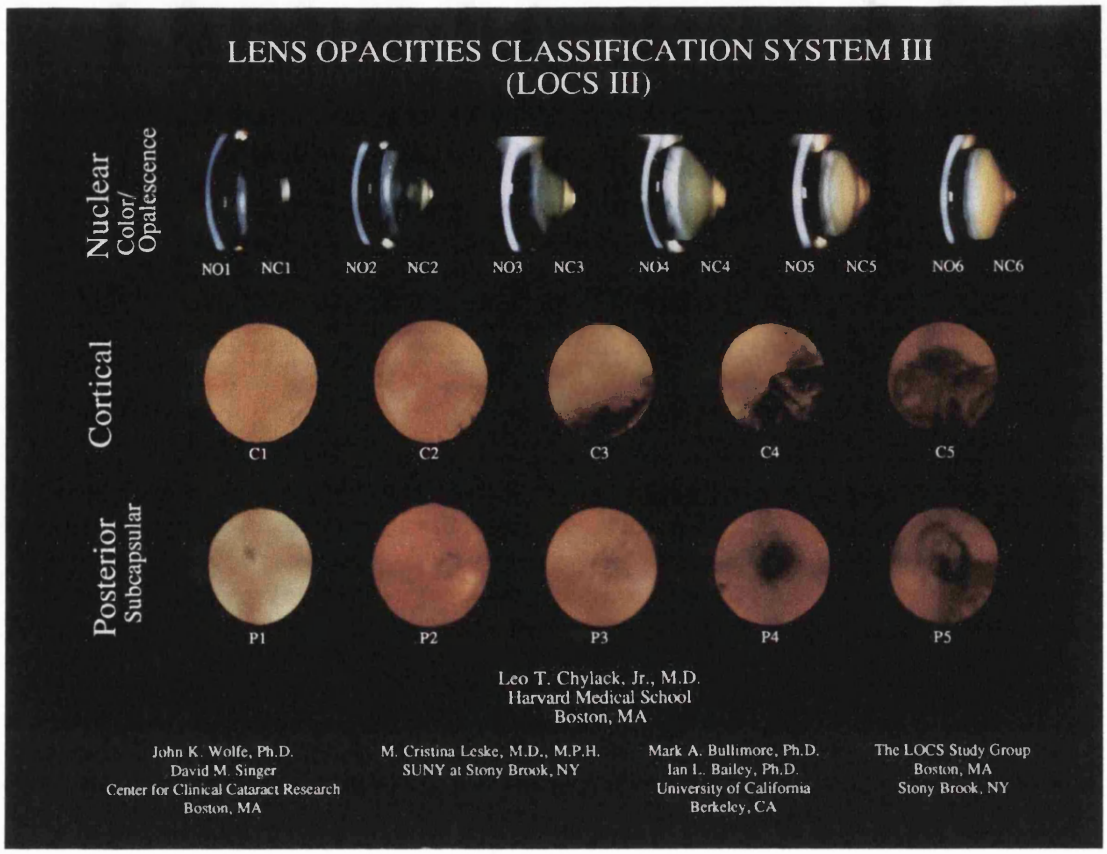


Figure 12 Lens Opacities Classification System III

By comparing the lens of a patient with the chart shown above it is possible to grade the following 4 variables in a lens: nuclear colour, nuclear opalescence, cortical cataract and posterior subcapsular opacity. Not shown are the pictures for a 0 rating for any of these variables. Furthermore the system is designed so that one can interpolate between the pictures shown and award grades to a precision of 0.1. Thus 2 of the 4 variables can be graded from 0-5, and 2 can be graded from 0-6, in 0.1 steps.

2.3 Visual Field Testing

Preoperative and postoperative automated perimetry was performed on a Mk II Humphrey Field Analyzer (HFA). The HFA was chosen because it is widely available and is the most frequently used automated perimeter in the UK (table 1). The decision to use the 24-2 program, as opposed to the 30-2 program was made because the latter takes longer and would create a few more "rim defects". All tests were performed on the same single machine except when the machine was not functioning when a type I machine was used (less than 10 visual field tests). Room lights were extinguished for all the tests, and a supervisor was present. The global field indices for all baseline and postoperative fields were entered into the trial database.

Prior to performing the visual field test, a standardised refraction was performed to determine the patient's distance refractive error and their best corrected visual acuity. Appropriate dioptric correction was used when the field test was performed (see below).

2.3.1 Preoperative Visual Field Testing

Where a patient had low reliability indices (see 2.1.1– Inclusion Criteria) that would exclude them from the trial it was possible to ask the patient to return for additional field tests. In order to minimise any learning effect seen in visual field testing all patients had to have completed at least two 24-2 fields prior to surgery. If a patient showed a "learning effect" as evidenced by a decrease in the mean defect of >2dB between the 2 baseline fields then he/she was asked to return for more visual fields until 2 successive fields show an improvement in mean defect of <2dB. Only the final baseline field was used in any subsequent analyses.

2.3.2 Postoperative Visual Field Testing

Post operatively field tests were obtained at 3 monthly intervals for the first year and then at 4 monthly intervals thereafter. Additional field tests could be requested where clinically indicated.

2.3.3 Refractive Correction for visual field testing

Prior to performing the visual fields test, a standardised refraction was performed. The distance refraction and patient details were entered into the HFA. The analyser contains software to calculate the appropriate refractive correction based on the patient's age and distance refraction.

2.3.4 Pupil Size and field testing

The patient's pupil size was assessed on the slit lamp using the lowest illumination settings possible. If the diameter of the pupil of the trial eye was <2.0mm, one drop of tropicamide 1% was instilled. This cut-off point was based on work showing that the

decrease in mean defect due to diminishing pupil size is greater in those with a pupil diameter of 2 mm or less (Webster et al. 1993). Only when maximum mydriatic effect was achieved was the test proceeded with. Other cycloplegics could be used if necessary. If a mydriatic was used then the patient was be re-refracted for near, prior to field testing, to see if any latent hypermetropia had been exposed.

2.4 Analysis of Visual Field Data

The initial field dataset available for analysis consisted of 56 patients followed over 16 months/ 6 fields. In order to provide a comparison against new change criteria developed in this thesis field change within this dataset was analysed using techniques already widely available:

a) Statpac global indices

b) Change criteria based on the Statpac II Glaucoma Change Probability software.

c) The AGIS grading system

The same field dataset was then analysed using a variety of pointwise linear regression criteria. Several PLR change criteria with the potential for more general application became apparent. These "potential" criteria were then applied to a) the same dataset followed for an additional 2 fields and b) a separate group of subjects from the same trial followed using the same protocol over 6 fields.

2.4.1 Statpac Global Field Indices

The visual field tests analysed in this work were performed on a single HFA MkII machine. Statpac global field indices were entered on a pro forma during each visit. The data was then entered in to the trial database. Using the Kolmogorov-Smirnov test the global field values were shown to be normally distributed. Values at 16 months were compared with those at baseline using a paired t test. The MD and PSD values were also analysed using linear regression. Although the Statpac software does perform a linear

regression of MD, the data was calculated directly because the software will under certain conditions omit one of the 2 baseline fields when performing linear regression of MD.

2.4.2 Statpac II Glaucoma Change Probability Analysis

For STATPAC II analysis the data was transferred using 5 ¼ inch floppy discs (3M, USA) to a type I HFA machine containing a copy of the STATPAC II software. STATPAC II Glaucoma Change Probability Change printouts were obtained for each patient. The printouts used the last baseline and first post-op tests as baseline fields for the Statpac analysis. Glaucoma change analysis was performed for each post trabeculectomy field test up to and including the 16 month visit. Two analyses were performed using the glaucoma change analysis printouts. For each patient the following were recorded:

a) The number of points showing significant progression or improvement at the 16 month stage. Patients were labelled as showing change if they had a greater number of progressing points compared to improving ones or vice versa.

b) The number of points showing significant change in the same direction in all 3 fields obtained at the 9, 12, and 16 month visits (fields 4, 5, and 6). This was done to demonstrate consistent change and to attempt to remove change detection due to noise.

2.4.3 AGIS grading

In order to provide an established reference point the fields were graded using an existing grading system. The AGIS scores for the initial 56 fields at baseline and at 16 months were calculated manually, according to the directions in AGIS study (AGIS 1994). A change in score of 4 or more was deemed to represent a significant change, as in the original study. The AGIS system was chosen in preference to that used by the Collaborative Normal Tension Glaucoma trial because the fields obtained in the trial could be directly analysed without any modification. For example the cNTG study requires 3

baseline fields and progression must be verified using 2 sets of 3 fields done within 1 month and 3 months later (Collaborative Normal-Tension Glaucoma Study Group 1998).

2.4.4 Pointwise Linear Regression Analysis (PLR)

The Humphrey Visual Field Analyzer does not contain software capable of performing PLR analysis. Data for such an analysis was therefore transferred to a personal computer using the same 5 ¼ inch floppies as those used for the Statpac analysis. Using an early version of the Progressor Software the HFA data was converted into a text data file. The text file contained data for each test including: name, hospital number, data and time of test, test duration, test type, foveal threshold, threshold values for each point tested. Where a point was tested twice the mean of the 2 thresholds was recorded. The text file was then imported into an Access 97 (Microsoft Corporation, Seattle, USA) database. Into this database tables from the main MRC 5-FU trial database were also imported. Within these imported tables are each patient's clinical details such as date of surgery, intraocular pressure, and global visual field indices. By constructing a query within this database it was possible to generate an array of point thresholds and test dates. The query result was then linked to a Microsoft Excel file. By linking the files it was possible to obtain a refreshed list of data in the spreadsheet every time that additional field data was added to the database. Excel's Visual Basic for Applications was used to write software to perform pointwise analysis of the visual fields. The code and sample Excel files are contained in a CD supplied with the thesis.

When performing pointwise linear regression over a field series, individual points were classified as progressing, improving, or showing no change. A progressing point is one whose slope is less i.e. more negative than the specified cut-off point and whose statistical significance is less than the specified level. An improving point is one whose slope is greater than the specified cut-off point and whose significance is less than the specified level. All other points were deemed to be showing no change. Each time a patient's fields

were analysed using any criterion the numbers of progressing and improving points were determined. Furthermore if a patient had both progressing and improving points then it was recorded as showing "Both". The Visual Basic software was written so that it would analyse field series from as many patients as desired at the same time. The software was run in 2 stages. The first stage used linear regression to calculate the slope and P value for each point for each patient over a specified number of fields (for the majority of this thesis this was 6). In the second stage the software counted the number of points improving or progressing in each series for individual change criteria. This 2 stage process meant that many criteria could be tested without having to recalculate the slope and P value for each point. Both the number and spatial arrangement of points progressing/ improving could be specified in a change criterion. Large numbers of criteria could be tested by holding either the P value constant and varying the slope or by holding the critical slope constant and varying the P value.

All patients who had been completely followed up for 16 months post-operatively as of 28/2/98 were included in the initial analysis. Patients were not analysed if a field test had been missed through failing to attend clinic. The aim of the analysis was to derive progression criteria from a set of visual fields where all patients had undergone the same frequency of visual field testing and the same duration of follow up. By setting these limitations 56 eyes from 56 patients had complete field sets that were available for analysis.

2.4.5 PLR Change Criteria

The initial change criteria to be assessed required a minimum of 1, 2, 3, 4, and 5 points significantly changing anywhere in the field. No restriction was made on the relationship between points. Either the critical slope was held constant as the P value was altered or the slope was altered as the P value was held constant. This made it possible to assess the effect of varying a single parameter on detection of change. As with the pointwise analysis performed using Statpac II it was frequently noted that a large number of patients could be

classified as progressing and improving in the same analysis. Thus again patients were labelled as both if they showed both progression and improvement.

2.4.5.1 Basic analysis

The last preoperative field and 5 post operative fields were analysed. The slope was held constant at 1 dB/yr while the critical significance/ P value was varied in 0.001 steps between $P<0.001$ and $P<0.05$. The number of patients showing 1, 2, 3, 4, or 5 points changing (progressing, improving, or both) at each P value was calculated. A value of 1 dB/yr was selected because this represents approximately a rate of change 10x greater than that due to age.

2.4.5.2 Slope analysis

The P value was held at either $P<0.05$ or $P<0.01$ while the critical slope was continuously varied in 0.1 dB/yr steps between 0.1 and 5 dB/yr. Again the number of patients showing 1, 2, 3, 4, or 5 points changing (progressing, improving, or both) at each slope value was calculated.

All subsequent work was performed with the P value varied within the range 0.001 – 0.05; subsequent analyses showed that within the ranges described above the change in P value had a more profound effect on numbers of patients changing than did the change in critical slope. No additional information was obtained from graphing numbers of patients changing against slope.

From the initial analyses it became clear that lax progression criteria were probably labelling patients as changing whereas in fact they were only detecting noise. Stricter criteria were developed in an attempt to reduce noise detection. Attempts were made to spatially relate changing points, temporally relate points, and to deal with points with excessive variability.

2.4.5.3 Spatial Change Criteria

Criteria were developed to preferentially select significantly changing points that had a spatial relationship. For a patient to be labelled as changing the points had to i) contiguous

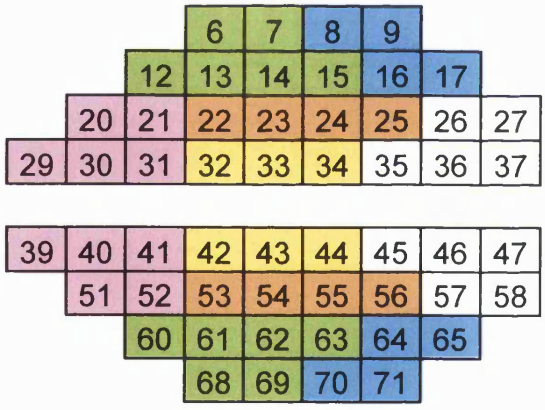
or ii) sited in a cluster. The clusters were derived by previous workers who had divided the visual field up into anatomically related areas. Finally a further attempt at spatial relationships was made by applying a Gaussian spatial filter to each field to assess its effect on detection of change.

The following change criteria were applied in addition to performing the basic analysis:

a) Contiguous points: a minimum of i) 2 and ii) 3 contiguous points had to be significantly changing in the same direction for a field series to be labelled as changing. Points could not be contiguous across the horizontal midline.

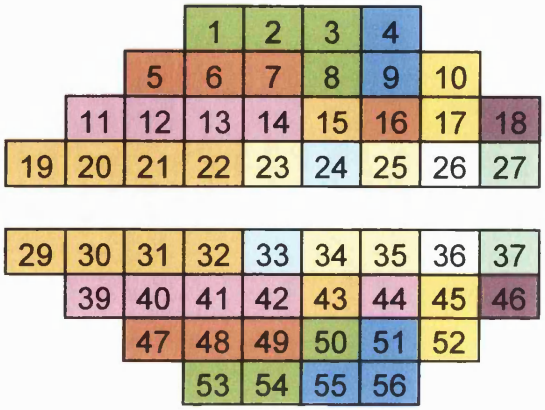
b) Glaucoma Hemifield Test Cluster: The number of eyes showing 2 points changing in the same direction in the same Glaucoma Hemifield Test Cluster (Åsman and Heijl 1992). The arrangement of these clusters is shown in figure 13.

Figure 13 Clusters in the Glaucoma Hemifield Test and Perimetric Nerve Fibre



Glaucoma Hemifield Test

The numbers refer to the order that the points are referred to in the HFA's computer output file. Some numbers are missing because they refer to points in other field test patterns(Åsman and Heijl 1992; Anton, Yamagishi et al. 1998)



Perimetric Nerve Fibre Bundle Map

Bundle Map.

c) Perimetric nerve fibre bundle cluster (PNFB): 2 points changing in any (PNFB) cluster. Figure 13 shows the arrangement of points in clusters based on Weber's work on the relationship between perimetric nerve bundles and visual field loss (Weber et al. 1990; Weber and Ulrich 1991). The actual cluster arrangement for the 24-2 field is taken from a subsequent study that mapped focal optic disc damage to visual field changes using a SLO (Anton et al. 1998). Some of the cluster areas as shown in fig 3 only cover a single point when a 24-2 field is considered. Given the expected low variability of the single cluster close to the fovea a significantly changing point in these clusters (numbers 24 and 33 in the diagram for the right eye) was sufficient to label a patient as changing. Elsewhere in the field 2 points had to be changing in the same cluster to label an eye as changing. The single clusters adjacent to the blind spot are likely to exhibit high variability being close to a physiological scotoma (Haefliger and Flammer 1989) and were therefore excluded.

d) The number of patients showing any 2 points changing anywhere in the area covered by all the GHT clusters that constitute each hemifield. This analysis was designed to see if any changes detected with a) were merely due to the effect that points around the blind spot are omitted by the distribution of the GHT clusters

e) The basic analysis was repeated after the individual fields had had a spatial filter applied to them. To filter the data the software considers each automated visual field as a grid of points and applies the following Gaussian filter to the edge points:

1	2	1
2	4	2
1	2	1

Filtering was not performed across the horizontal midline.

2.4.5.4 Temporal Change Criterion

The temporal relationship between points was assessed by repeating the basic analysis with the additional proviso that points were only labelled as changing if they had a significant slope when analysed over 5 and over 6 fields.

2.4.5.5 Excessive Point Variability

An alternative method of dealing with noise was to either remove from the analysis some of those points known to show greater variability or to apply stricter progression criteria. One analysis removed the points above and below the blind spot while a second technique looked at using stricter slope or P value criteria for more peripheral points.

a) Removal of Blind Spot Points: Previous workers have removed the points above and below the blind spot presumably because of the large amount of variability they show (Katz et al. 1997; Wild et al. 1997). Such points around this physiological scotoma show increased variability (Haefliger and Flammer 1989). The basic analysis was thus repeated excluding the points lying above and below the physiological blind spot.

b) Stricter Criteria For Edge Points: Previous work by our group (Viswanathan et al. 1996; Viswanathan et al. 1997) used progression criteria that set a critical slope for points at the edge of 30-2 fields that was double that of the inner ones. This was designed to reduce excessive detection of change in peripheral points due to increased peripheral noise. Here the basic analysis was repeated but with the critical slope for edge points doubled or the P value for the edge points halved.

2.4.5.6 Comparison of Change Across the Horizontal Midline

Any potential progression attributable to glaucoma suffers from the confounder of change due to cataract. Using the assumption that any cataractous change would produce a symmetrical effect on change in both vertical hemifields, further progression criteria were developed that compared progression across the horizontal midline. The simplest compared the absolute difference in the number of changing points between the 2 vertical

hemifields. The second compared the difference in the number of changing points between the symmetrical clusters defined in the glaucoma hemifield test (GHT). The difference in the number of points changing between the upper and lower hemifields was calculated. Again this was done over both 5 and 6 fields. This analysis was then refined so that the progression criterion became a difference of 2 changing points between mirror GHT clusters above and below the horizontal midline.

2.4.5.7 Random Analysis

In order to assess the effect of random noise in the analysis the basic analysis was repeated with the order of the threshold values individually randomised for each point. In order to smooth out random fluctuations the randomisation was performed 50 times for each field series. The results of the basic analysis were then divided by 50 to obtain the mean random change.

2.4.5.8 Mean Slope and Threshold of Changing Points

The mean slope and baseline threshold of all changing points over both 5 and 6 fields was calculated as the P value was changed in 0.001 steps between $P<0.001$ and $P<0.05$. Mean slopes and baseline thresholds were also calculated for the random analysis performed in 2.4.1.8.

2.4.6 Comparison of Progressors and Non-progressors

For each PLR criterion that yielded exclusive progression a list of patients progressing and not progressing was generated. Progressors and non-progressors were compared using non field measurements. The measurements included: pre- and post-operative IOP, global Statpac indices, foveal threshold, LOCS scores, and visual acuities. The small numbers of patients progressing (4-6 out of 56) coupled with the large number of comparisons prevented any statistical expression of significance. However the median values of these parameters were graphically compared used box-whisker plots. Since LOCS scores were not available for the 16 month visit, an attempt was made to see if there was any

association between logMAR visual acuity or foveal threshold and LOCS grade. This was in the hope of finding an indirect marker for field changes attributable to cataract development.

2.5 Validation of PLR progression criteria

Having examined the distribution of numbers of eyes showing change over 6 fields, favourable change criteria were selected. They were then applied to 2 additional data sets to see if the changes detected using them were repeatable. The 2 data sets were:

a) Fields acquired from patients entering the trial after the initial 56. Only those patients who had undergone at least 6 fields were analysed. This mimics the process used in developing neural networks where results from a training set (criteria derived from the original 56 eyes) are then applied to a test set (those patients joining after the initial 56).

b) The 6 fields from the original 56 patients studied above plus a subsequent 2 fields, providing a total analysis over 8 fields. This provided an opportunity to see if similar trends in change were demonstrable over longer follow up.

Only those criteria that continued to show exclusive progression using the first data set were applied to the additional data sets. By examining the detection of progression and improvement by criteria across the 3 data sets it was possible to empirically choose criteria that would be appropriate for future use.

3 Results

3.1 Follow-up of patients within trial

Recruitment for the MRC 5-FU filtration surgery study began on 1/4/96. As of 28/2/98 when I stopped working on the trial 183 patients had agreed to take part, 172 had undergone a trabeculectomy.

Table 8 summarises the recruitment and follow up status at this point. One years complete visual field data was available for 60 patients; the figure for 16 months was 32. By using data collected after 28/2/98 it was possible to analyse data on 56 patients at the 5 and 6 field stage (12 and 16 months post-op respectively).

Table 8 MRC 5-FU trial recruitment and patient progress as of 28/2/98.

Number of patients:	
Agreeing to take part in trial	183
Undergone trabeculectomy	172
Lost to follow up	1
With 16 months of complete post-op data	56

Table 9 summarises the baseline age, gender, and racial distribution of those 56 patients. More than 56 patients have been followed up to the 6 field time point. The information on these patients has not been analysed because it is incomplete, mainly because the patients missed visits due to sickness, or because the patients were unable to complete an HFA test due to intercurrent ocular conditions. A few patients were not included because equipment failure had resulted in field tests not being stored on the HFA. It was possible to manually re-enter the threshold values from a printout for linear regression analysis but not possible to create an HFA file to permit Statpac analysis.

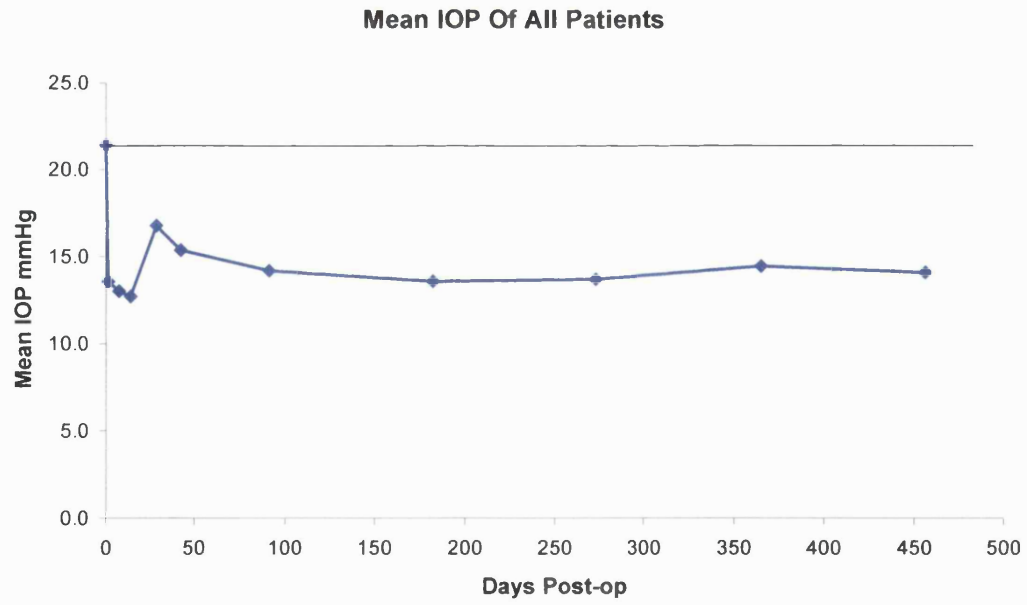
Table 9 Baseline characteristics of the 56 trial patients in the initial analysis.

Age	Mean 66 yrs (Range 45-80)
Gender	18F : 38M
Race	Caucasian 52, Black 2, Asian 2
Mean months since glaucoma diagnosed	57.6 (Range 0-281)
Mean Deviation (dB)	-10.4 (Range –0.5 to –29.8)

3.1.1 IOP changes following trabeculectomy

The mean IOP versus follow up is shown in figure 14 respectively. Mean IOP at pre assessment (not listing) was 21.4 mmHg the mean values at 12 and 16 months were 14.1 and 14.2 mmHg respectively (t test $P<0.0001$ for both visits). The mean IOP at listing was 23.9 mmHg. The change in IOP between listing and pre assessment was significant ($P<0.01$ t test).

Figure 14 Mean IOP in the MRC 5-FU following trabeculectomy.



Mean IOP falls immediately following trabeculectomy. It rises over the next 6 weeks to reach a plateau where it stabilises for the next 15 months.

3.1.2 Visual Acuity and Foveal Threshold Following Trabeculectomy

The mean logMAR near and distance visual acuities and foveal thresholds at baseline, 12 and 16 months post-op are summarised below (table 10), distribution plots of

the change in distance VA and foveal threshold are shown below in figure 16. There was no significant change in visual acuity or foveal threshold compared to baseline at either 12 or 16 months.

Table 10 Mean visual acuity and foveal threshold at baseline, 12 and 16 months.

	Baseline	12 Months	16 Months
Distance logMAR VA	0.06	0.09	0.09
Near logMAR VA	0.23	0.27	0.30
Foveal Threshold	32.2	-0.4	-0.6

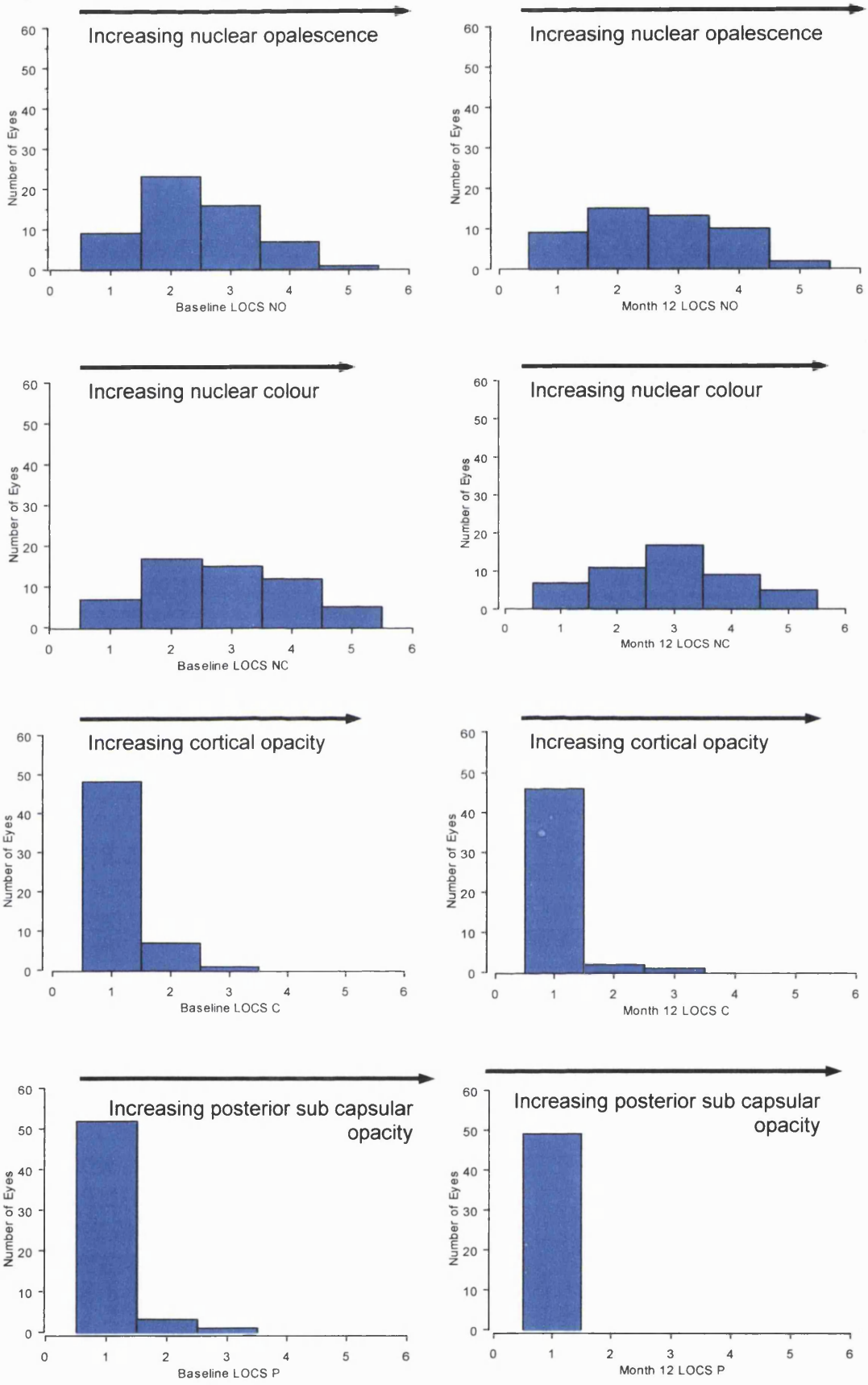
None of the values at 12 or 16 months show a significant difference (t test) compared to its respective baseline value

3.1.3 Cataract Grading

Distribution plots of modified LOCS grading scores at baseline and 12 months post-operatively are shown in figure 15. The change in the distributions of nuclear colour (NC) grades yielded a P value=0.039 (Chi-Squared). There was no significant change in the nuclear opalescence (NO) grades. The high frequency of grade 1 scores for cortical and posterior subcapsular lens opacity precluded any analysis of change of distribution of these gradings. It does appear as though there was a reduction of posterior subcapsular cataract (P), preoperatively several patients had grades 2-3, while everyone had grade 1 post operatively.

Using the Spearman Rank Correlation test there was no evidence of any association between any modified LOCS grading and baseline foveal threshold. Nor was there any correlation between foveal threshold and cortical or PSCLO grades at 12 months. There was a significant association between foveal threshold and NO ($r=-0.51$, $P<0.001$) and NC ($r=-0.50$, $P<0.001$) at 12 months. There was no correlation between change in modified LOCS grade and change in foveal threshold over 12 months.

Figure 15 Modified LOCS grading at Baseline and 12 Months



There is a small but significant rightward shift in the distribution of nuclear colour (NC). No significant change in the distribution of nuclear opalescence is detected. The distribution of cortical (C) and posterior sub capsular (P) opacities is so biased towards grade 1 that analysis of their distribution is not possible.

There was no significant correlation between distance visual acuity (VA) and modified LOCS grade at baseline. Distance VA was significantly associated with NO ($r=0.44$, $P<0.001$) and NC ($r=0.51$, $P<0.001$) at 12 months, but not with Cortical or PSCLO. There was no association between near visual acuity and any modified LOCS grade at baseline or 12 months.

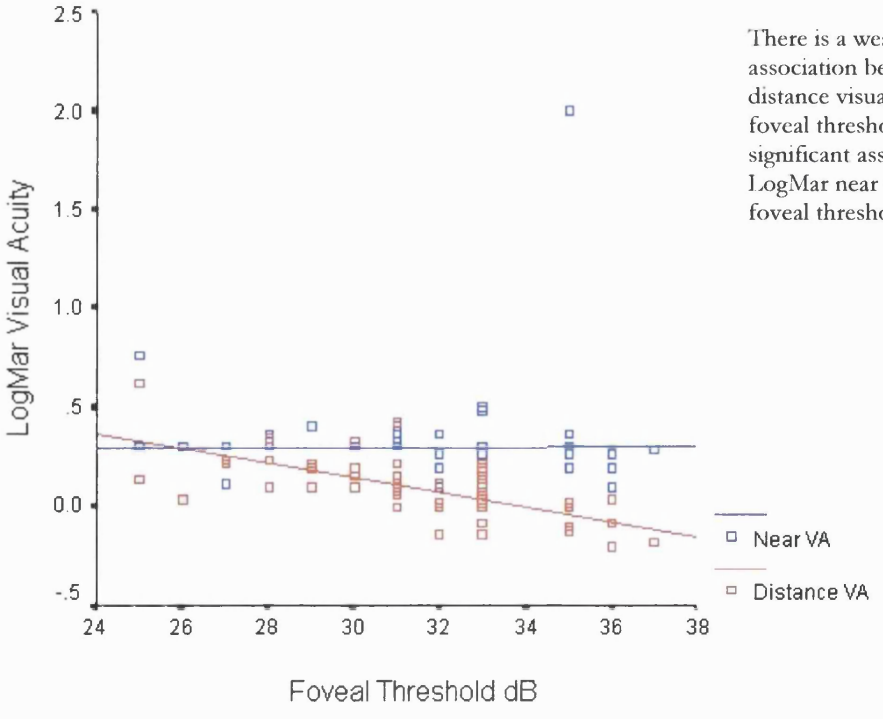
Using the Pearson Correlation Foveal threshold (FovTh) was found to be significantly associated with distance visual acuity at baseline, and at 16 months.

Table 11 Correlation Between Foveal Threshold and Best Corrected logMAR Visual Acuity

		logMAR Visual Acuity	
		Baseline Distance	**Baseline Near**
Baseline Foveal Threshold	Pearson Correlation	-0.39	-0.06
	Significance	0.0028	0.6586
		16/12 Distance	**16/12 Near**
16/12 Foveal Threshold	Pearson Correlation	-0.68	0.01
	Significance	<0.0001	0.9496

If one makes a Bonferroni correction to allow for multiple analyses then only distance logMAR visual acuity is significantly associated with foveal threshold.

Figure 16 Near and Distance LogMar VA Versus Foveal Threshold at 16 Months.



There is a weak but significant association between LogMar distance visual acuity and foveal threshold. There is no significant association between LogMar near visual acuity and foveal threshold.

Distance but not near visual acuity is significantly associated with foveal threshold at 16 months. It was not associated with near visual acuity at baseline or 16 months (table 11). It should be borne in mind that a significant association does not imply clinical significance.

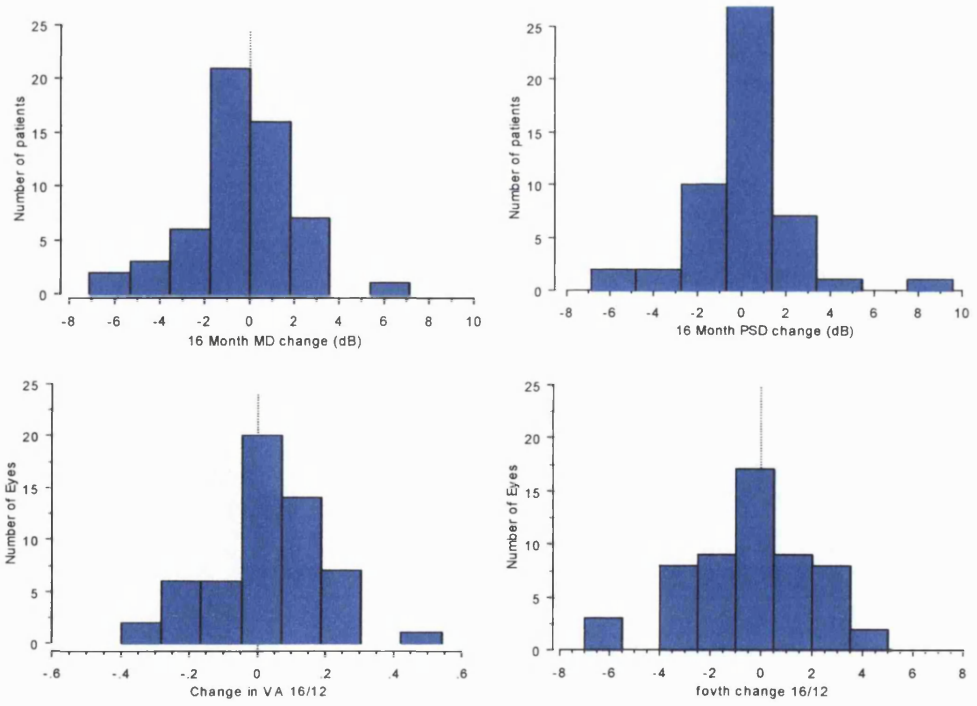
Figure 16 shows the near and distance VA plotted against Foveal Threshold at 16 months. The association for distance VA is significant while the association for near VA is not.

3.2 Statpac Analysis

3.2.1 Changes in STATPAC global indices

Values for MD, PSD, SF, and PSD at baseline, 3,6,9,12, and 16 months are shown below (table 12). None of the changes at 16 months was significant when compared using a t test with the baseline values. The distributions of the changes in MD, PSD, logMAR distance VA, and foveal threshold over the first 16 months are shown below (figure 17).

Figure 17 Distribution of change in MD, PSD, logMAR distance VA and foveal threshold at 16 months



The changes in the 4 parameters between baseline and 16 months are evenly distributed around zero, suggesting little real change.

Table 12 Statpac Global Indices for each post operative visit from baseline – 16 months.

	Pre op	Month 3	Month 6	Month 9	Month 12	Month 16
MD	-10.40	-9.73	-10.76	-10.39	-10.38	-10.80
MD (SD)	8.15	9.18	7.84	8.06	7.95	7.87
SF	2.37	1.97	2.12	2.31	2.36	2.58
SF (SD)	0.98	0.97	1.02	1.23	1.11	1.38
PSD	7.38	7.44	7.37	7.40	7.44	7.47
PSD (SD)	3.82	4.00	3.89	3.94	3.91	3.71
CPSD	6.55	7.05	6.77	6.82	6.79	6.93
CPSD (SD)	4.02	4.05	4.00	3.99	4.11	3.92

None of the changes in MD, SF, PSD, and CPSD were significant at 16 months when compared to baseline.

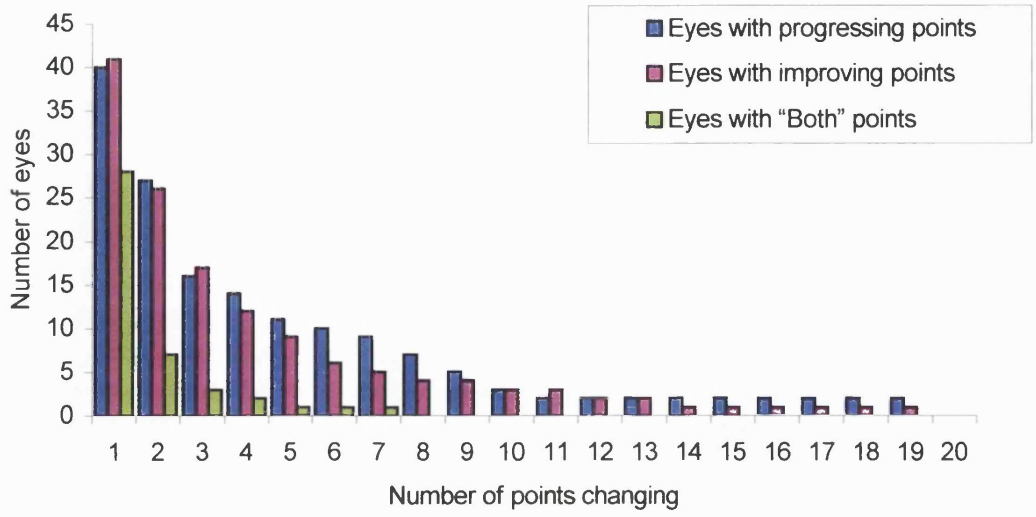
3.2.2 Linear regression analysis of global indices

Linear regression of the mean MD and PSD values from the 56 patients yielded slopes of –0.41 dB/yr and 0.05 dB/yr respectively. Neither slope was significant.

3.2.3 Pointwise analysis using Glaucoma Change Probability analysis

The number of field series containing points progressing, improving or both at 16 months is shown below (figure 18). It is clear that a large number of patients showed a large number of changing points. Furthermore progression and improvement were detected in almost equal amounts. This meant that there was no obvious minimum number of changing points criterion that would exclusively select any more than one progressing patient. The same information is also displayed in figure 38 where a patient is labelled as changing if there are more points changing in one direction than another.

Figure 18 Whole field Glaucoma Change Probability analysis at 16 months.



Graph showing the number of eyes changing against "the number of points changing"-the minimum number of individual points required to be changing in the same direction using GCP analysis in order to label the eye as changing. Approximately the same number of eyes contain equal numbers of points showing progression or improvement. Large numbers of improving or progressing points are seen in 3 eyes (2 progressing and 1 improving).

If one uses a stricter change criterion that requires a point to be consecutively changing in the same direction at the 9, 12 and 16 month visits (fields 4, 5, and 6) less overall change is detected. Eight patients show improvement while only 5 show progression (figure 38)

3.2.4 AGIS scoring

The AGIS scores for individual patients are shown in figure 38. The scores at baseline and at 16 months show how little change is detected using an existing criterion. Five patients show progression while 2 show improvement.

3.2.5 Correlation of global indices with IOP

Pearson Correlation coefficients were used to compare the change in global Statpac indices and the change in IOP over the first 16 postoperative months (table 13). Change in IOP and change in global indices were shown to be normally distributed using the Kolmogorov-Smirnov test. A level of significance was deemed to be $P<0.0125$ since 4

analyses were performed. No significant association was seen between change in IOP and change in global indices.

Table 13 Pearson correlation between change in IOP and change in global Statpac values at 16 months.

	MD Change	PSD Change	SF Change	CPSD Change
IOP Change	0.284	0.091	0.225	0.073
Sig. (2-tailed)	NS	NS	NS	NS

No significant association between global Statpac values and the change in IOP over the first 16 postoperative months is seen.

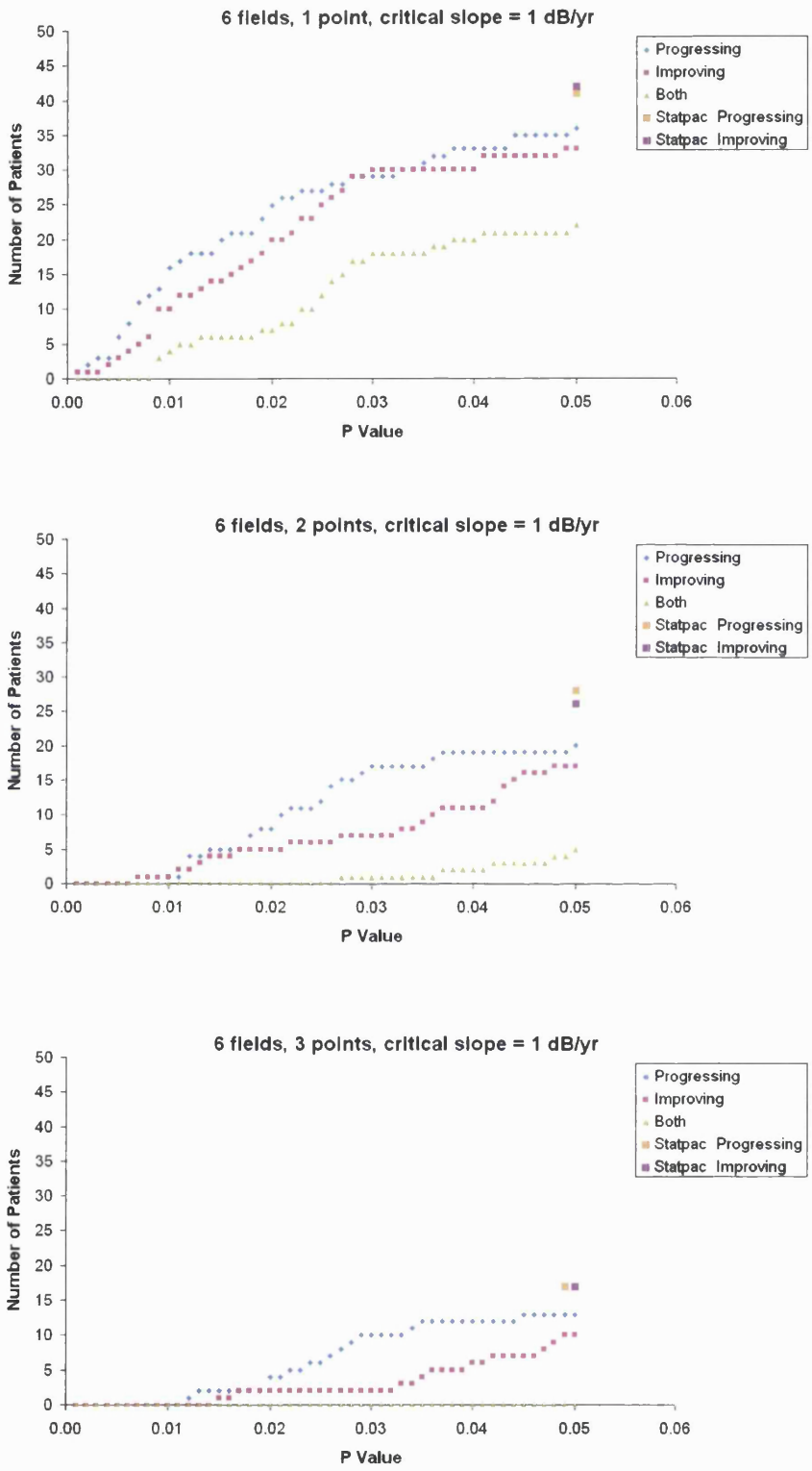
3.3 Pointwise linear regression

The results in this section deal with varying criteria for detecting change. Fifty six field series from 56 patients were analysed over 6 fields, representing 16 months of post-operative follow up. In the following section a series is labelled as having changed after linear regression analysis of all the points in a series. If a specified number of points show a significant slope with specified minimum slope and maximum significance levels then the whole field series is deemed to be changing. Change may involve a reduction in sensitivity, "Progression", or an "Improvement". Furthermore within the same field series some points may show improvement while others show progression. In such a scenario a series may be labelled as "Both".

3.3.1 Basic Analysis

The basic analysis is an exploration of the criteria developed in previous studies by our group (Fitzke et al. 1996; McNaught et al. 1996; Bhandari et al. 1997; Viswanathan et al. 1997) and also the initial criterion selected for the MRC 5-FU filtration surgery study trial – 1 point with a slope of 1 dB/yr and $p<0.05$. Figure 19 shows the change over a 6 field series in numbers of eyes showing progression, improvement or both when the significance level was varied between 0.001 and 0.05 with the critical slope held at 1 dB/yr. Charts are shown for a 1, 2, 3, 4 or 5 point change

Figure 19 Numbers of eyes showing change against P value. 6 fields



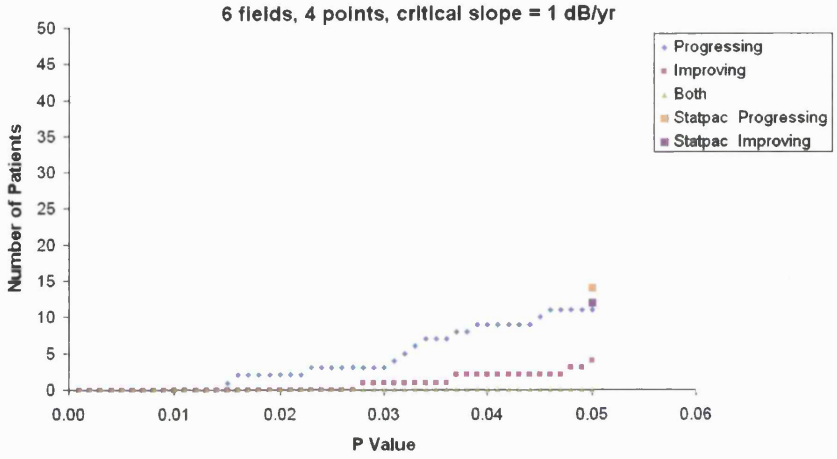
Large amounts of progression, improvement and "both" are detected at lax P values. As a stricter P value is used detection of change decreases. No exclusive progression is seen At P<0.05 much more change is seen using PLR than Statpac.

Less change is seen compared to using a single point criterion. At middle range P values more progression than improvement is detected. No exclusive progression is seen.

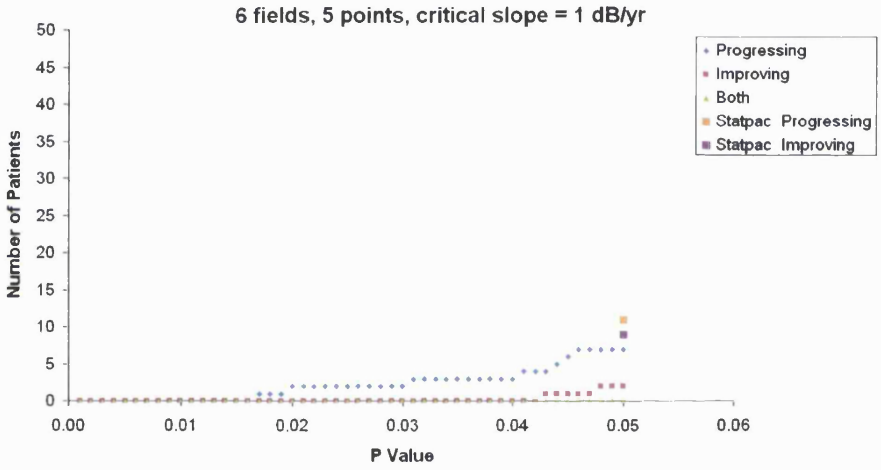
Overall the amount of change detected continues to fall. At P<0.05 the amounts of change detected using PLR is less than that detected using Statpac. At no P value are both progression and improvement seen in the same eye.

1, 2, or 3 points, Slope= 1 dB/yr. Also shown are the numbers of eyes changing using the Statpac Glaucoma Change Probability analysis for same number of points.

Figure 19 Continued. Numbers of eyes showing change against P value. 6 fields, 4 or 5 points, Slope= 1 dB/yr.



The amount of change continues to decline with 3 eyes showing detection of exclusive progression at P<0.027



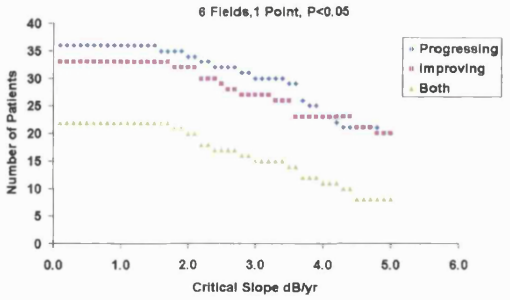
Detection of improvement is almost abolished with exclusive detection of progression in 4 eyes at P<0.042. Progression and improvement continue to be detected in almost equal amounts using Statpac

criterion. Also shown are the number of eyes who would be labelled as progressing or improving if the change criterion required the same number of points to be changing on the Statpac Glaucoma Change Probability analysis.

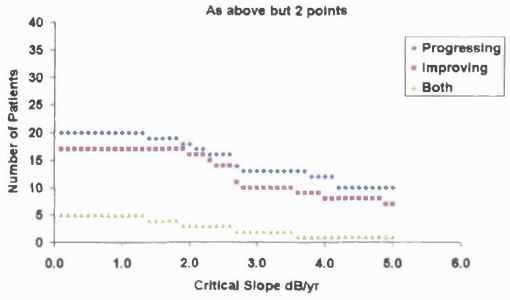
3.3.2 Slope Analysis

The slope analysis is a repetition of the basic analysis, with the level of significance held constant at $p<0.05$ while the critical slope is varied. Figure 20 shows the change over 16 months in the numbers of patients showing progression, improvement or both when the critical slope was varied between 0.001 and 5.0 dB/yr with the significance level held at 0.05. Charts are shown for a 1, 2, 3, 4 or 5 point change criterion.

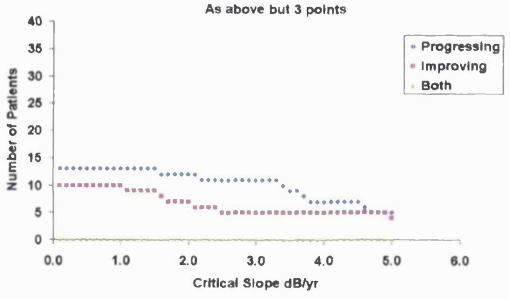
Figure 20 No. of eyes showing change v slope. 6 fields, 1-5 points, P< 0.05.



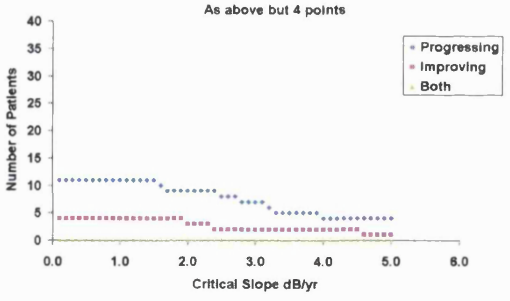
Large amounts of change are detected a low slope values. Increasing critical slope from 0.01 to 1.5 dB/yr has no impact on the amount of change detected. Even using a large critical slope figure still leads to large amounts of change detected



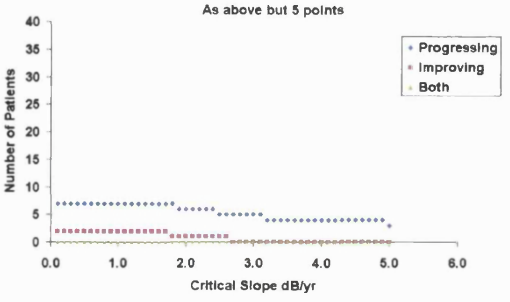
Less change is detected compared to using a single point but a similar trend is observed. "Both" is detected for all critical slope values used.



With a 3 point criterion no "both" is detected. More than 10 eyes show either progression or improvement between 0.1 and 1.0 dB/yr. No exclusive progression is detected.



There is a preferential reduction in detection of improvement rather than progression compared to using 3 points. Still no exclusive progression is detected



Exclusive progression is detected for values of critical slope 2.7 dB?yr and above.

Overall the effect of changing the critical slope from 0 to 5 dB/yr is less profound than changing the level of significance from P<0.001 to 0.05. Changing the slope from 0 to 1 dB/yr has no effect in any of the graphs

Whether change was plotted against critical slope or P value the amount of change detected fell as the slope became greater or the P value became stricter. With very lax criteria e.g. using 1 point changing over 6 fields with a slope of 1 dB/yr and a P value <0.05, a large amount of change was detected, 33 patients were classed as improving, 36 progressing and 22 showed both improvement and progression (figure 19 and figure 20). Equally with the use of stricter change criteria all detection of change could be abolished. No patients are labelled as changing using a criterion of 3 points with a slope of 1 dB/yr and a P value<0.011 (figure 19). No patient shows both progression and improvement if the minimum criterion requires that 3 or more points have a significant slope (figure 19 and figure 20) regardless of P value.

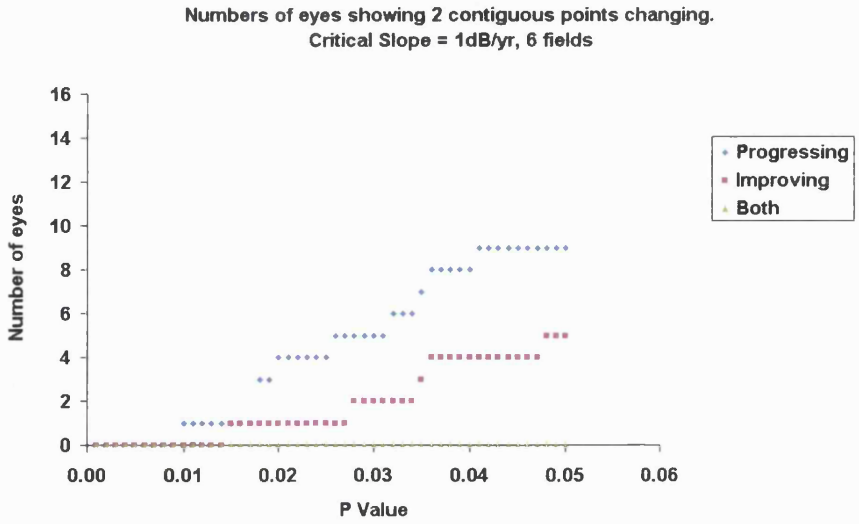
Over 6 fields the maximum number of patients that could be exclusively detected, using the above criteria, as progressing was 5 using a criterion of 5 points, critical slope of 2.7 dB/yr and P<0.05 (figure 20).

3.3.3 Spatial Change Criteria

The preceding criteria only stipulated that a given number of points had a significant slope anywhere within the visual field. The next step was to examine the effect of employing a stricter criterion for change by requiring points changing in the same direction to be spatially related.

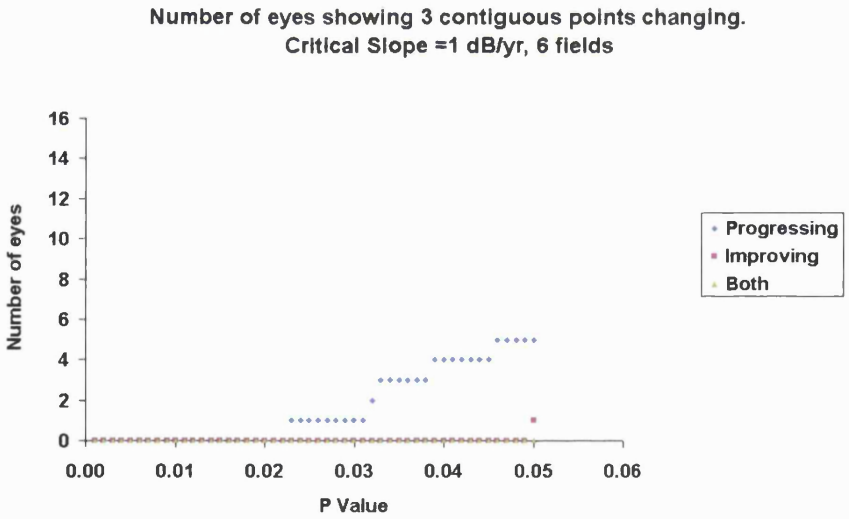
Figure 21 and figure 22 show the numbers of field series showing change when the minimum requirement for improvement or progression is that 2 (Figure 21) or 3 (figure 22) contiguous points in the same upper or lower half of the field are changing. The amount of change detected is reduced compared to that seen in the basic analysis (figure 19) where 2 or 3 points anywhere in the field are required to be changing. A 3 contiguous point criterion detects less change than a 2 contiguous point criterion. Five patients could be labelled as exclusively progressing if a criterion of 3 contiguous points, with a P value <0.049 was used.

Figure 21 Change criterion - 2 contiguous points, 6 fields Slope = 1 dB/yr. Number of eyes showing change against P value.



The amount of change detected is reduced compared to that seen using 2 points in the in the basic analysis. Only 1 eye would be detected as showing exclusive progression.

Figure 22 Change criterion - 3 contiguous points, 6 fields, Slope = 1 dB/yr. Number of eyes showing change against P value.

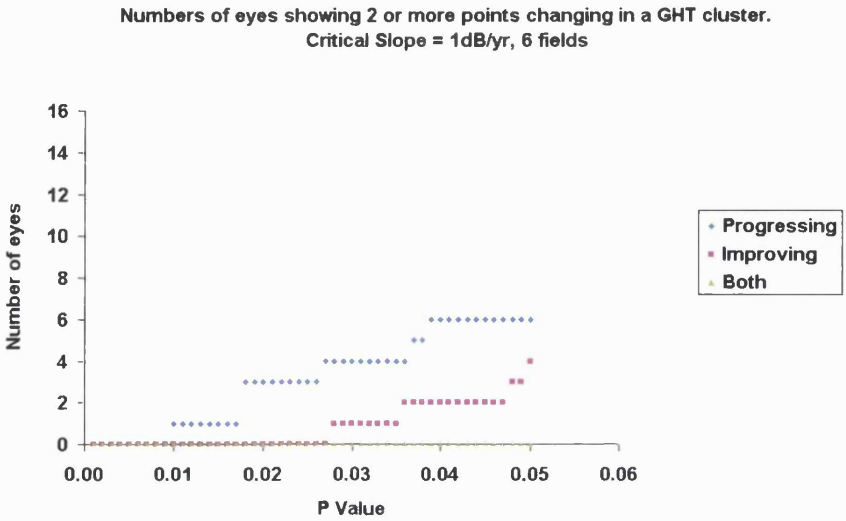


There is good separation of progression and improvement curves with 5 eyes showing exclusive progression if a P value <0.049 is used.

a) The amount of change detected when requiring 2 points to change in a glaucoma hemifield cluster (figure 23) is less compared to using a criterion of 2 points changing anywhere in the visual field (figure 19). An exclusive progression criterion of 2 points changing in a GHT cluster, P<0.027, slope=1 dB/yr can be derived. In all 4 patients

detected using this criterion the progressing points are contiguous although the criterion does not require them to be.

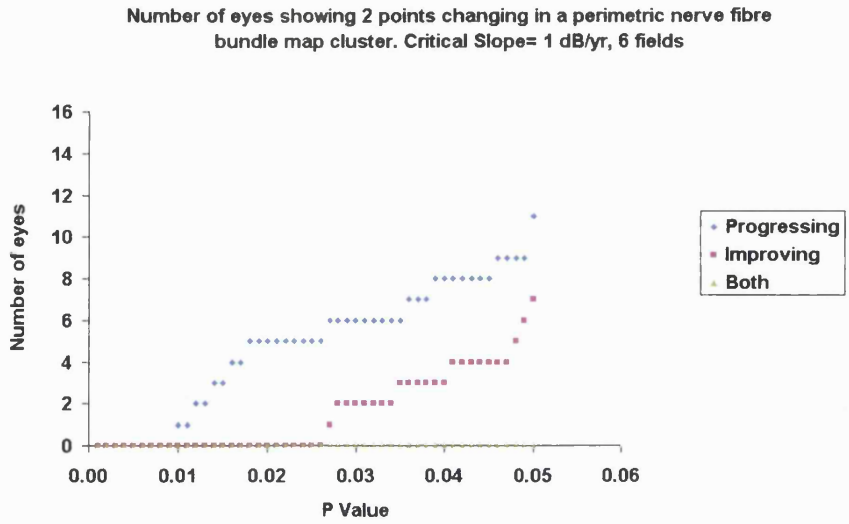
Figure 23 Change criterion - 2 points in a GHT cluster, Slope=1 dB/yr, 6 fields. Numbers of eyes showing change against P value.



Less change is seen that using the criterion of 2 contiguous points. However 4 eyes show exclusive progression at P<0.027.

b) Using the perimetric nerve fibre bundle PNFB map clusters detects slightly more change (figure 24) compared to using GHT clusters. There is a greater leftwards movement in the progression curve than the improving curve. By using a criterion of 2 points changing in a PNFB cluster (single point changing in foveal cluster), P<0.026, slope=1dB/yr, five patients are exclusively detected as progressing.

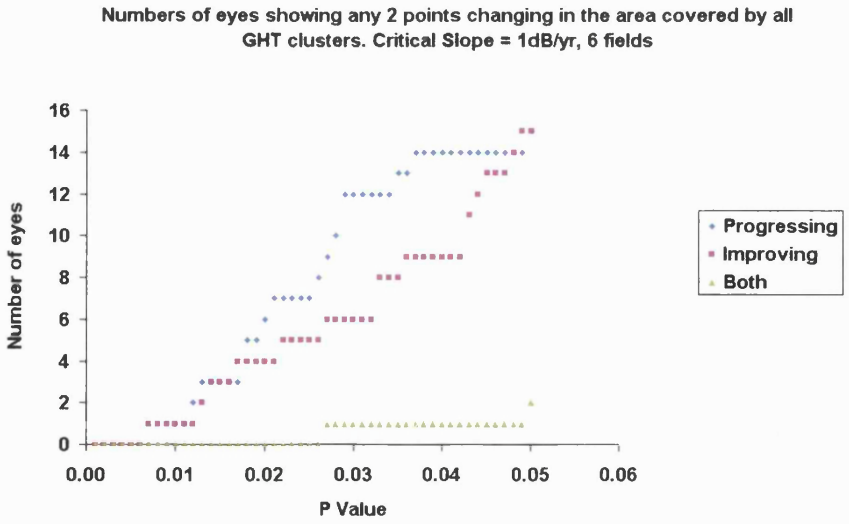
Figure 24 Change criterion - any 2 points in a perimetric nerve fibre bundle, 6 fields, Slope = 1 dB/yr. Numbers of eyes showing change against P value.



There is increasing separation of the 2 curves compared to figure 17. Five eyes show exclusive progression.

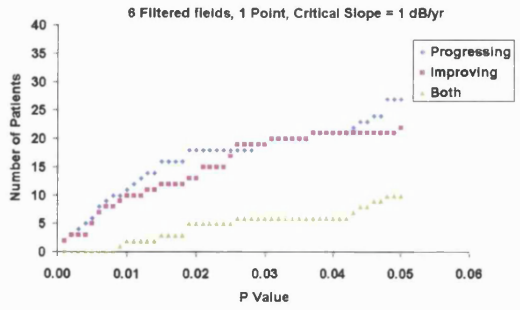
c) Where the change criterion is 2 points with significant slopes anywhere in the area of field covered by all GHT clusters rather than within individual clusters then the number of change detected increases dramatically (figure 25). Allowing for differences in vertical scales the graphs in figure 25 are similar to those in and figure 19 albeit slightly reduced.

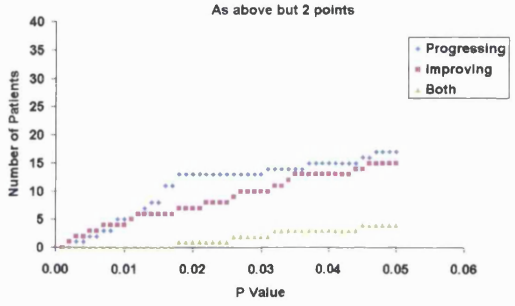
Figure 25 Change criterion - any 2 points in area covered by GHT, 6 fields, Slope = 1 dB/yr. Numbers of eyes showing change against P value.

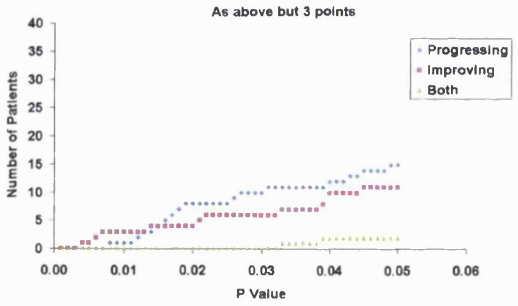


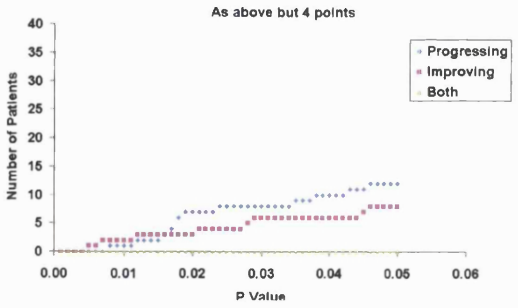
There is poor separation of the 2 curves with no detection of exclusive progression

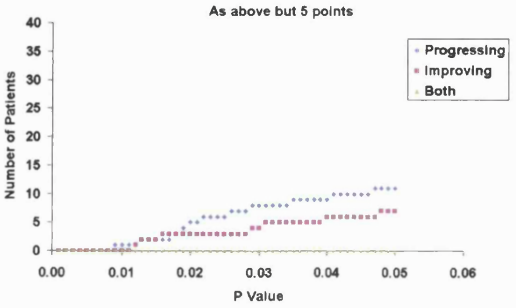
Figure 26 Change criterion – 1-5 points, 6 Filtered Fields. Slope=1 dB/yr. Numbers of eyes showing change against P value.











No obvious benefit is seen from applying the filter. Levels of change are similar compared to figure 19. The detection of exclusive progression seen with 4 or 5 points in the basic analysis is abolished after applying the Gaussian filter.

d) Following the application of a simple Gaussian filter to the data from each field test the numbers of patients showing change against P value was plotted in figure 26. No clear benefit is seen from applying the filter. In fact when using a 4 or 5 point change criterion the use of the filter prevents the exclusive detection of progression. Furthermore examining figure 26 shows that with a 3 point criterion some patients show both improvement and progression whereas none show both when using unfiltered fields (figure 19).

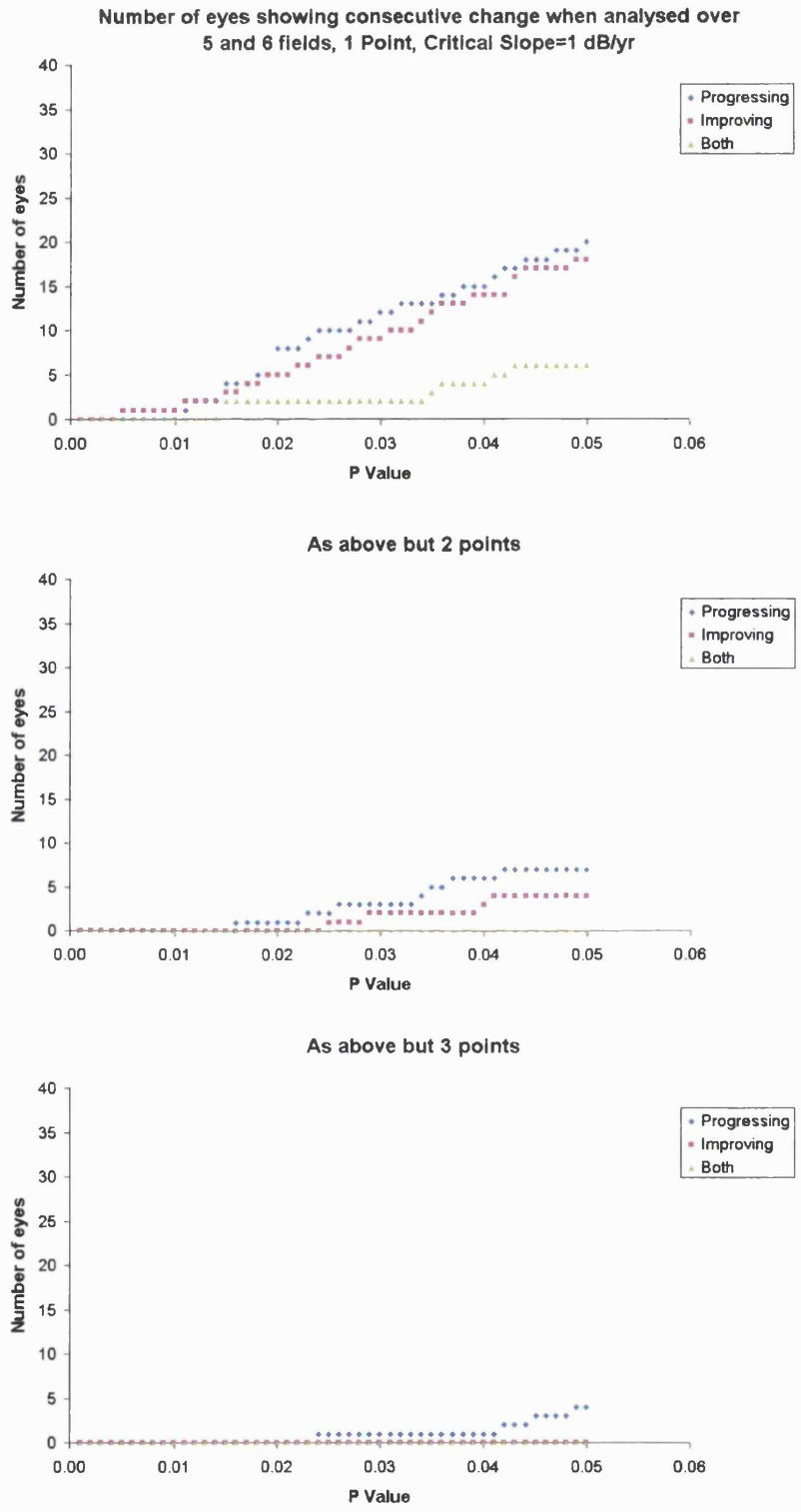
3.3.4 Temporal Change Criterion

Figure 27 shows the number of patients showing change at successive analyses over 5 and 6 fields. In order for a series to be classified as showing change, the same progressing or improving point must show change in the 5 and 6 field series. In general the amount of change detected is reduced compared to that detected over 6 fields alone. Small amounts of exclusive progression are found using 2 point change criteria

3.3.5 Increased Point Variability

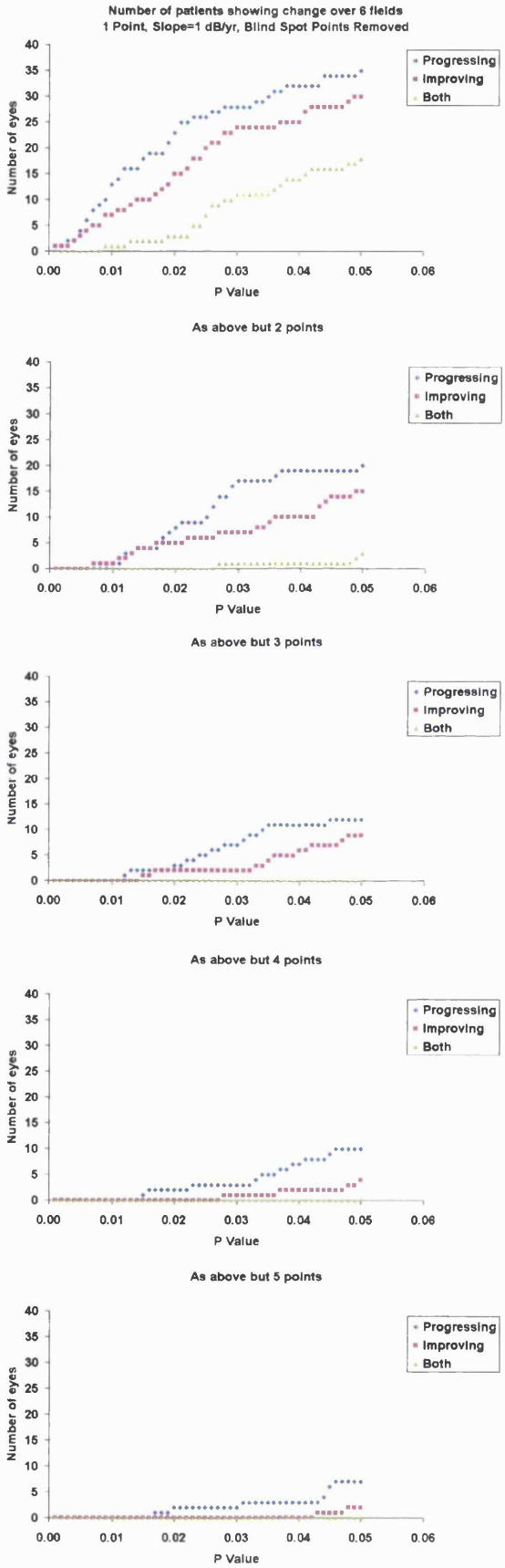
a) Removal of Blind Spot Points: Figure 28 shows the amount of change detected when the points above and below the blind spot are removed from the basic analysis. Very little effect is noticed when compared to the basic analysis in figure 19 except in the plots where a single point with a significant slope is used as the criterion for labelling the patient as changing. Under these conditions there is a slight reduction in detection of change, especially in the "improving" curve. This change does not help in the selection of an alternative exclusive progression criterion.

Figure 27 Numbers of eyes showing change at same point at 5 and 6 fields. 1, 2, or 3 points, Slope=1 dB/yr.



The amount of change detected is reduced compared to that detected over 6 fields alone. Small amounts of exclusive progression are found using 2 point change criterion. (Slope=1 dB/yr, P<0.024). Four patients show exclusive progression with an alternative 3 point change criterion (slope=1 dB/yr, P<0.050).

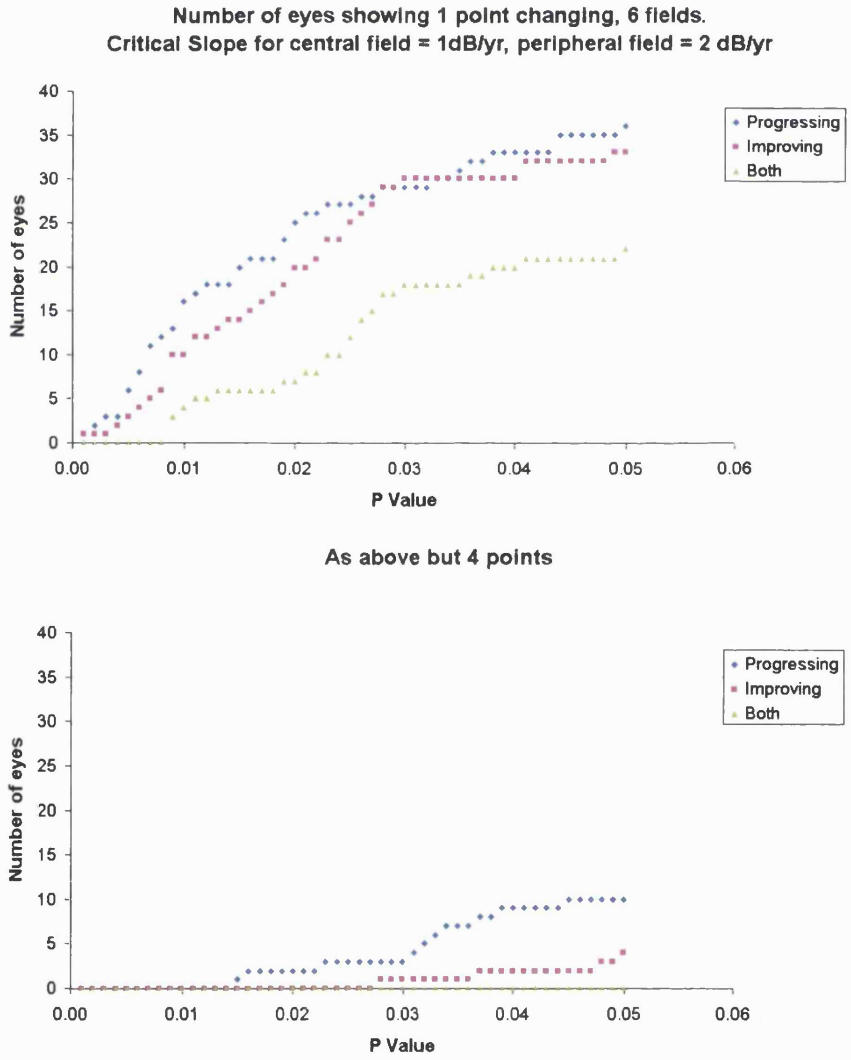
Figure 28 Change criterion – 6 fields, blind spot points removed, 1-5 points, Slope=1dB/yr. Change against varying P value.



Very little difference is noted compared to the basic analysis figure 19

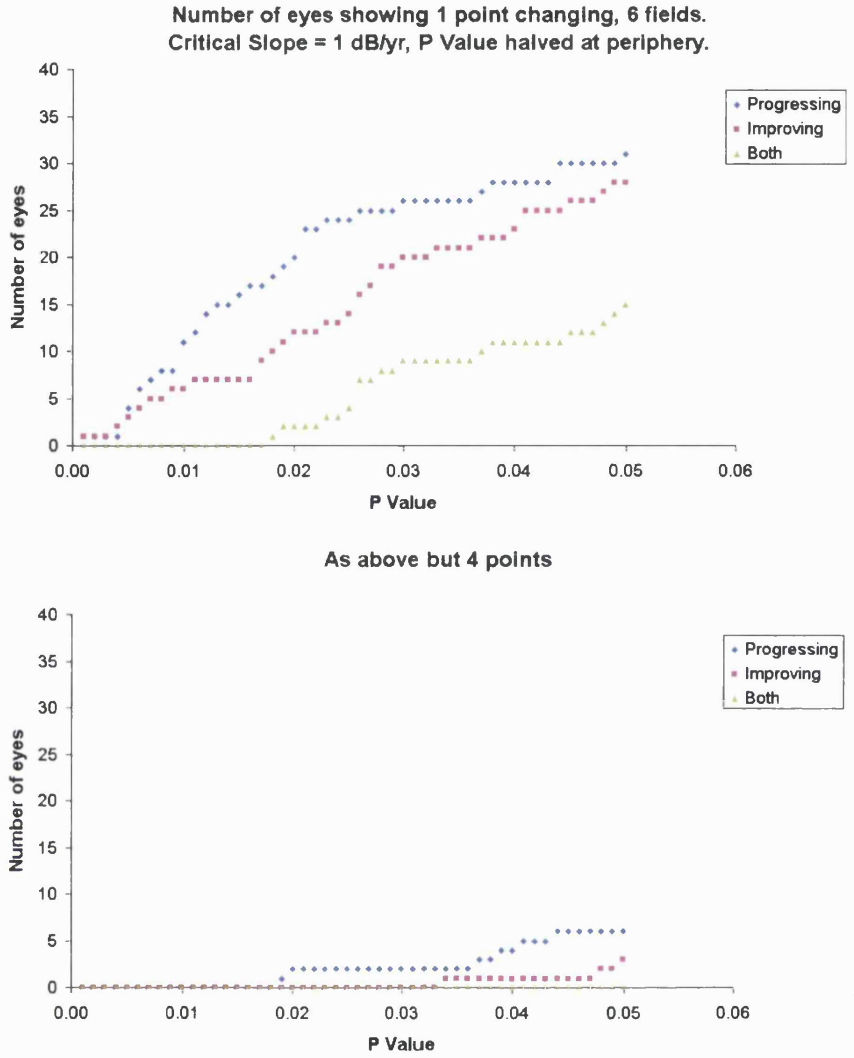
b) Stricter Criteria for Edge Points: Figure 29 shows the effect of using a critical slope for edge points that is double that of the remaining inner points. No obvious benefit compared to the basic analysis is seen whether single or 4 point change criteria are employed. Figure 30 shows the effect of halving the P value for edge points. Little effect compared to the basic analysis is seen with a single point criterion. With a 4 point criterion there is an overall reduction in the detection of change compared to the basic analysis. Doubling the slope or halving the P value for edge points did not increase the detection of exclusive progression.

Figure 29 Number of eyes showing change when the Critical Slope is doubled for peripheral points. 1 or 4 points. 6 fields, Critical slope centrally = 1 dB/yr.



No obvious benefit compared to the basic analysis is seen whether single or 4 point change criteria are employed.

Figure 30 Number of eyes showing change when the P value is halved for peripheral points. 1 or 4 points. 6 fields, P value centrally <0.05.

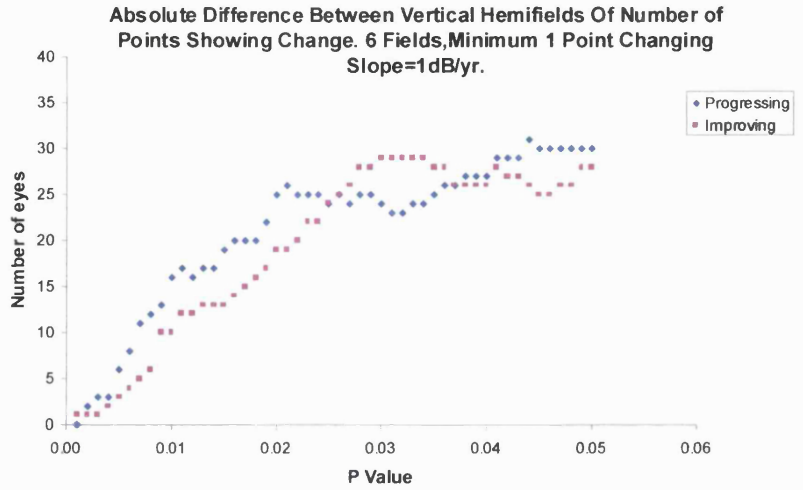


Little effect compared to the basic analysis is seen with a single point criterion. With a 4 point criterion there is an overall reduction in the detection of change compared to the basic analysis.

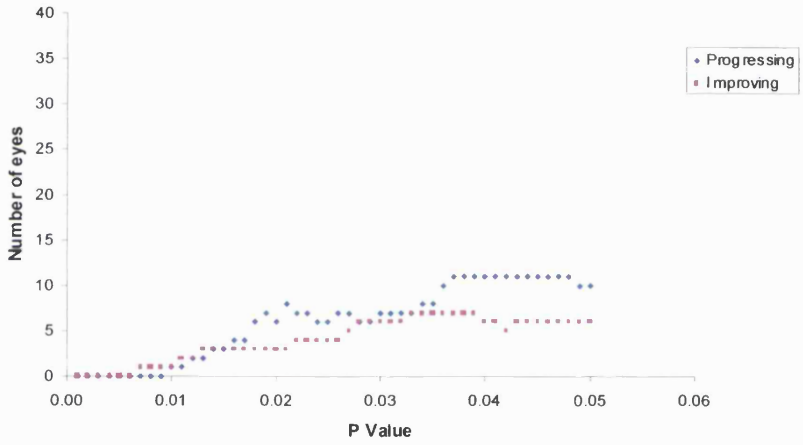
3.3.6 Comparison of Change Across the Horizontal Midline

Figure 31 show the numbers of patients labelled as changing when the change criterion is a specified difference in the number of significantly changing points between the two vertical hemifields. Where a single point difference is required the numbers of patients showing change are similar to those in the basic analysis (figure 19). Visual inspection of the data shows that this can be attributed to an asymmetric distribution of changing points across the horizontal midline.

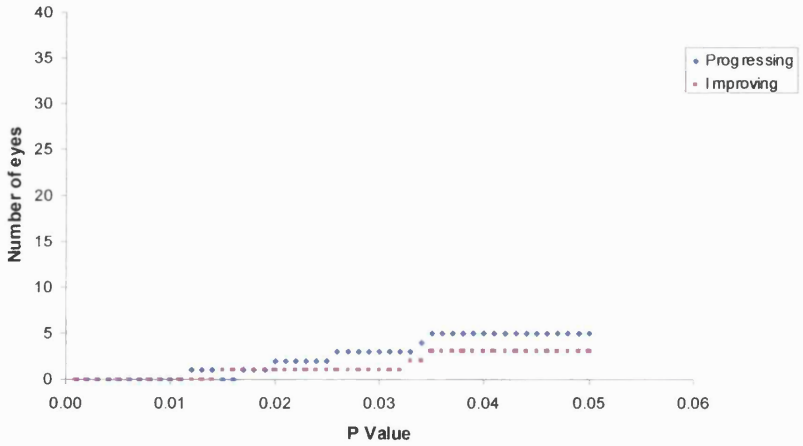
Figure 31 Change criterion - Difference of 1, 2, or 3 changing points between vertical hemifields. 6 fields, Slope =1 dB/yr. Change versus P value.



As above but minimum of 2 point difference

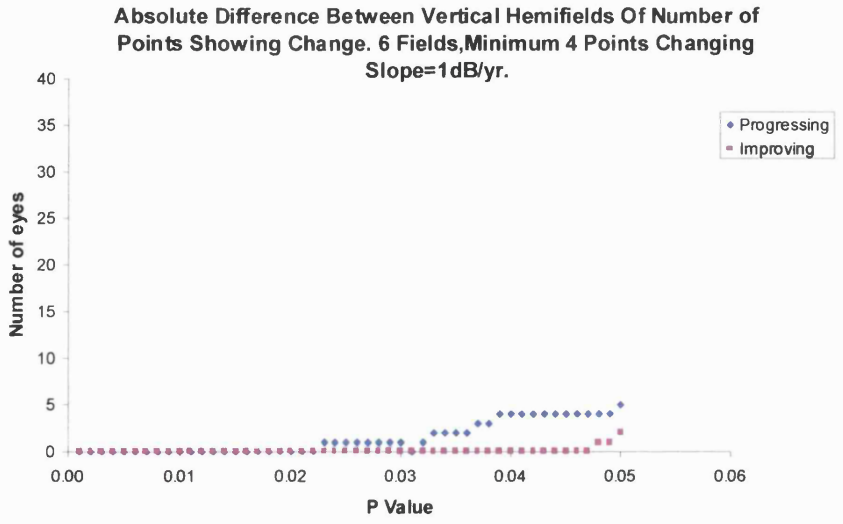


As above but minimum of 3 point difference

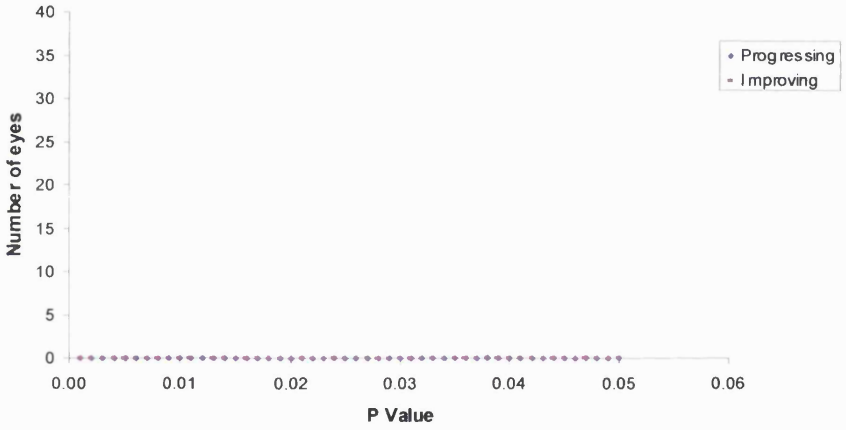


Where a single point difference is required the numbers of patients showing change are similar to those in the basic analysis.

Figure 26 Continued. Change criterion - Difference of 4 or 5 changing points between vertical hemifields. 6 fields, Slope =1 dB/yr. Change versus P value.



As above but minimum of 5 point difference

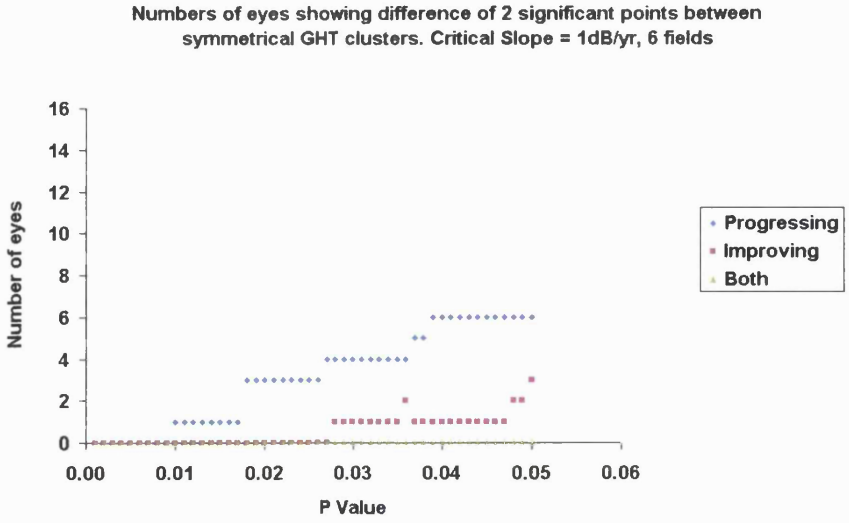


Requiring 2 or more points to change results in less change compared to the basic analysis. With a 4 point criterion 4 eyes show exclusive progression at P<0.039.

Requiring the difference in number of changing points to be greater than 1 reduces the overall amount of change detected. A criterion requiring at least 4 points to change with a P value <0.039 exclusively detects progression. There is a sudden reduction in detection of change when a 5 point difference is required with no change at all detectable.

The results of comparing mirror image GHT clusters and using a change criterion of a difference of 2 significant points between such clusters are shown in figure 32. The graphs are almost identical to those of figure 23. The same 4 patients show exclusive progression over 6 fields when a P value of <0.027 is used.

Figure 32 Change criterion - Difference between 2 symmetrical GHT clusters of at least 2 significant points, 6 fields, Slope=1 dB/yr.

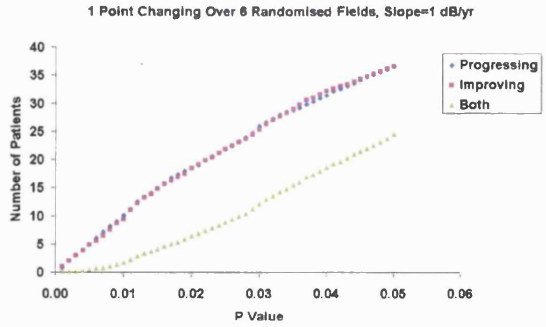


The graphs are almost identical to those in figure 23 that uses the change criterion of 2 points in a GHT cluster. No benefit is seen from comparing clusters across the horizontal midline. The transient increase in the "Improving" curve at $P<0.036$ arises because the plot derives from the differences between symmetrical clusters.

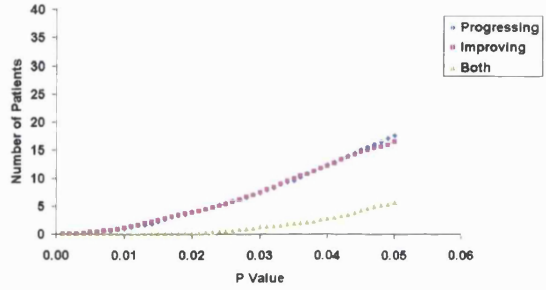
3.3.7 Random Analysis

Randomly rearranging the test sequence for each point would be expected to reveal the amount of detected change that could be attributed to random noise in automated visual field testing. The numbers of patients showing change in field series with randomly rearranged test sequence is shown in and figure 33. There is little difference in the graphs of actual change (figure 19) and random change when 1, 2, or 3 point change criteria are used. As such it suggests that much of the change detected with such criteria is due to random change. It is only with stricter criteria requiring 4 or 5 points to change that one sees a marked separation between the random and actual graphs (figure 19).

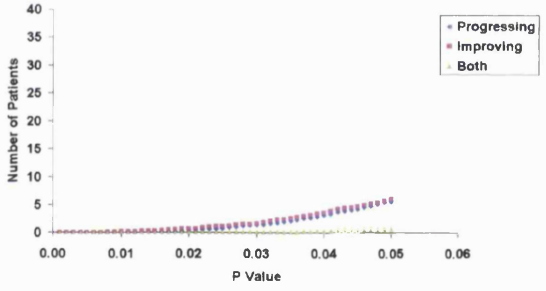
Figure 33 Number of eyes showing change against P value, 6 randomised fields, Slope =1 dB/yr.



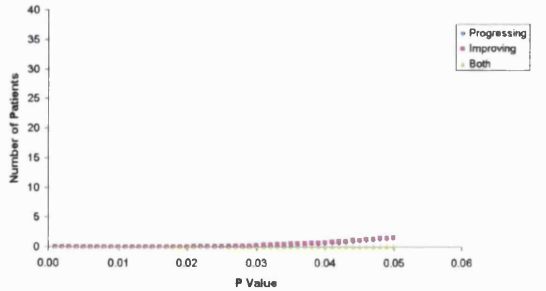
As above but 2 points



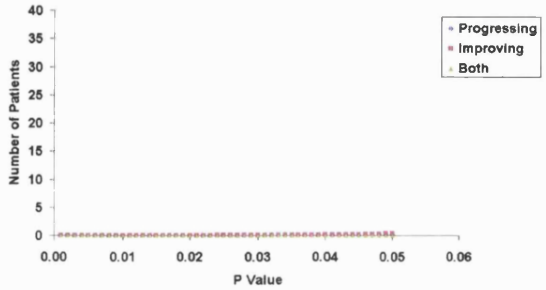
As above but 3 points



As above but 4 points



As above but 5 points



With 1-3 points large amounts of change are seen that are similar to those seen in the basic analysis in figure 13. This suggests that the change seen may be random.

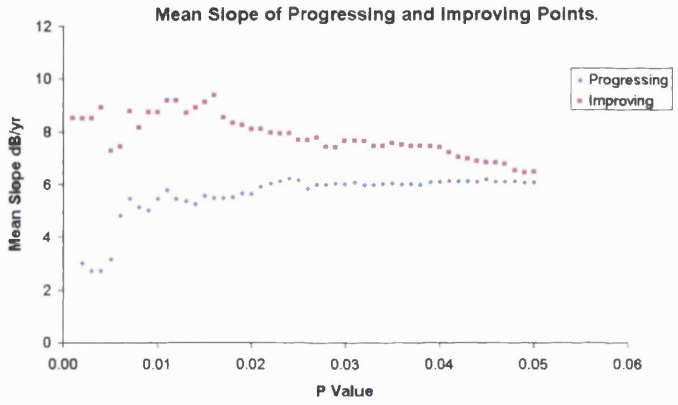
With 4-5 points the change detected in the randomised fields is less than the actual change detected in figure 13

3.3.8 Mean Slope and Baseline Threshold of Changing Points

The mean slope of progressing and improving points was plotted across a range of significance levels (figure 34). Any significant point was included regardless of slope. This analysis was repeated after the test order had been randomised (figure 35).

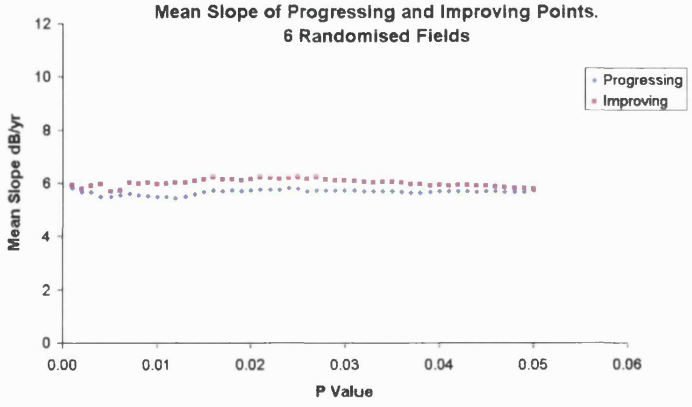
The mean baseline sensitivity was plotted for across a range of significance levels (figure 36). Any significant point was included regardless of slope. This analysis was repeated after the test order had been randomised (figure 37).

Figure 34 Mean slope of changing points versus P value. No minimum critical slope is required.



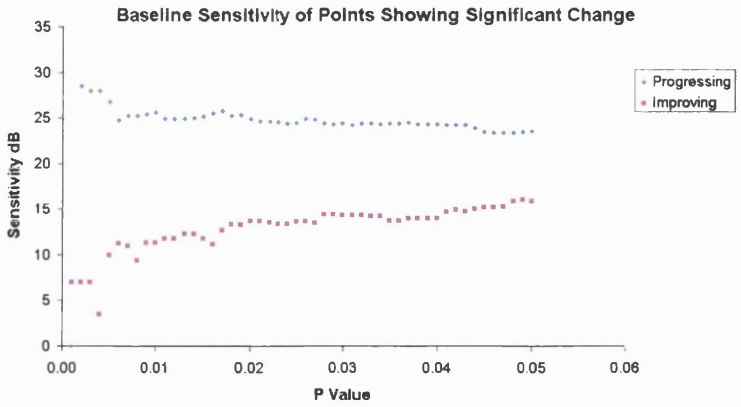
At strict P Values the mean slope of the improving points is lower than that of the progressing ones. However the slopes converge to around 6 dB/yr as the P Value becomes laxer.

Figure 35 Mean slope of significantly changing points in randomised field series. No minimum critical slope is required



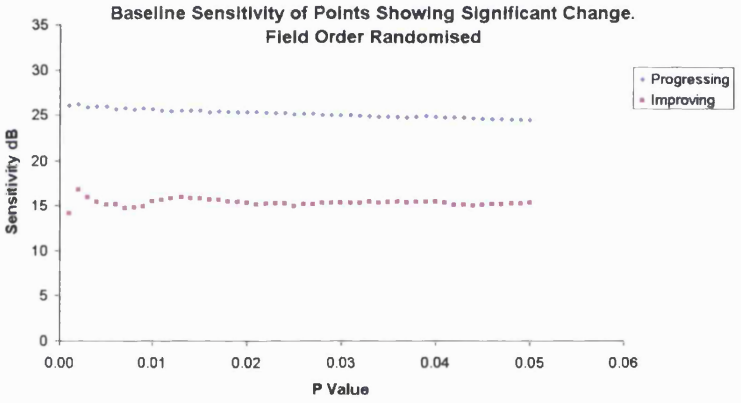
Changing the level of significance does not change the slope of the points detected if the field order is randomised. The mean slope for both curves is the same as that detected in the non randomised series at the P<0.05 level (figure 28)

Figure 36 Mean baseline sensitivity of points changing over 6 fields. No minimal slope is required.



The baseline sensitivity of progressing points is much higher than that of the improving points at strict P Values. There is some drift towards each other as laxer P Values are applied.

Figure 37 Mean baseline sensitivity of points changing over 6 randomised fields. No minimal slope is required



Very little impact on baseline sensitivity is seen as the P Value is changed. Both curves stay at approximately the same level that the corresponding curves in figure 30 reach at the P<0.05 level.

3.4 Progression criteria and change

Six criteria were derived that exclusively detected progression of 4 or more patients over 6 (or 5 and 6) fields. Figure 38 shows which patients were detected as changing using each criterion. Also shown in the figure those patients that could be deemed to be progressing by using criteria derived from the glaucoma change probability printout.

The pointwise linear regression criteria that detected exclusive progression were:

a) 5 points, critical slope=2.7 dB/yr, P<0.05. 5 eyes.

b) 3 contiguous points, critical slope=1 dB/yr, P<0.049. 5 eyes

c) 2 points in a GHT cluster, critical slope=1 dB/yr, P<0.027. 4 eyes

d) 2 points in a PNFB cluster, critical slope=1 dB/yr, P<0.026. 5 eyes.

e) 4 point difference in number of changing points between vertical hemifields, critical slope=1 dB/yr, P<0.047. 4 eyes.

f) 3 points found to be consecutively changing at 5 and 6 field analysis stage, critical slope=1 dB/yr, P<0.050. 4 eyes.

As can be seen from figure 38 there is some overlap of eyes detected as exclusively progressing by different criteria, however the overall range of eyes detected as exclusively progressing is large.

3.5 Progression criteria and media opacity

Foveal threshold and logMAR VA were used as surrogate markers for changes in media opacity. The small numbers of patients found to be exclusively progressing precluded any direct statistical comparison of mean changes in foveal threshold and logMAR visual acuity between those showing progression and those showing no change. However changes in logMAR visual acuity and in foveal threshold were compared graphically. Figure 40 and figure 41 show box-whisker plots for the median of the changes in foveal threshold and logMAR VA respectively for progressors and non-progressors as determined by the 6 exclusive progression criteria. The indentations in each box delimit the 95% confidence intervals for the median. Figure 39 explains the box-whisker plot in more detail.

Figure 38 Change shown in individual trial patients using PLR, Statpac II, Glaucoma change probability, and AGIS criteria

Trial No.	Baseline MD	Linear Regression Change Criteria - Progression Exclusively Detected						16 Month Statpac Analysis			Change over final 3 consecutive fields - Glaucoma change probability analysis		AGIS		
		6 fields 5 points P<0.05 slope=2.7	6 fields GHT cluster P<0.027	6 fields modified PNFB P<0.026	6 fields, difference between hemifields of 4 points P<0.039	Consecutive change at 5 & 6 fields, 3 points, P<0.050	6 fields 3 contig points, P<0.049,	No. of progressing points	No. of improving points	Progressing minus Improving	Number of progressing points	Number of improving points	Baseline AGIS score	6 Field AGIS score	Difference between 6 fields and baseline
23	-29.8	N	N	N	N	N	N	0	0	0	0	0	20	20	0
53	-27.8	N	N	N	N	N	N	0	2	-2	0	0	20	20	0
58	-26.1	N	N	N	N	N	N	0	4	-4	0	0	19	15	-4
65	-24.7	N	N	N	N	N	N	1	1	0	0	0	20	20	0
49	-24.1	N	N	N	N	N	N	1	1	0	0	0	19	16	-3
38	-24.0	P	N	N	P	N	P	8	0	8	3	0	16	20	4
34	-23.8	N	N	N	N	N	N	2	0	2	0	0	16	18	2
15	-21.9	N	P	N	N	P	N	1	1	0	0	0	17	17	0
20	-19.5	N	N	N	N	N	N	1	3	-2	0	0	14	12	-2
84	-18.4	P	N	N	P	N	N	3	1	2	1	0	13	16	3
5	-17.8	N	N	N	N	N	N	2	1	1	0	0	11	14	3
29	-17.3	N	N	N	N	N	N	1	2	-1	1	0	15	14	-1
31	-16.5	N	N	N	N	N	N	0	2	-2	0	0	14	13	-1
45	-16.0	N	N	N	N	N	N	1	1	0	0	0	13	12	-1
10	-15.6	N	N	N	N	N	N	0	2	-2	0	0	14	15	1
13	-15.5	P	N	N	N	P	P	2	1	1	0	1	13	17	4
30	-15.4	N	N	N	N	N	N	0	5	-5	0	1	11	9	-2
80	-14.8	N	N	N	N	N	N	0	5	-5	0	0	14	13	-1
24	-13.9	N	N	N	N	N	N	2	13	-11	0	2	13	10	-3
2	-13.4	N	N	N	N	N	N	8	7	1	0	0	9	12	3
28	-13.3	N	N	N	N	N	N	0	11	-11	0	0	14	8	-6
52	-13.2	N	N	N	N	N	N	0	6	-6	0	3	11	9	-2
25	-12.0	N	N	P	N	N	N	2	0	2	0	0	5	11	6

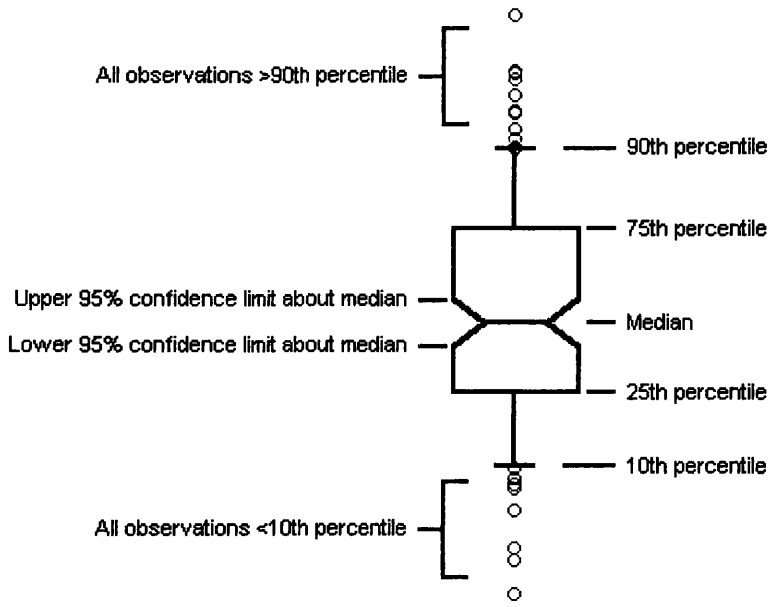
Exclusive progression: PLR citerion
Progression: non-PLR criterion
Improvement: non-PLR criterion

Figure 38 cont'd Change shown in individual trial patients using PLR, Statpac II, Glaucoma change probability, and AGIS criteria

25	-12.0	N	N	P	N	N	N	2	0	2	0	0	5	11	6
33	-10.7	N	N	N	N	N	N	7	0	7	0	0	7	9	2
44	-10.1	N	N	N	N	N	N	1	3	-2	0	3	10	9	-1
27	-7.8	N	N	N	N	N	N	0	1	-1	0	0	6	7	1
36	-7.6	N	N	N	N	N	N	0	2	-2	0	0	6	3	-3
46	-7.4	N	N	N	N	N	N	9	4	5	0	0	8	9	1
47	-7.4	N	N	N	N	N	N	0	0	0	0	0	7	10	3
68	-6.9	N	N	N	N	N	N	2	3	-1	0	0	7	6	-1
17	-6.8	N	N	N	N	N	N	0	9	-9	0	0	5	5	0
50	-6.0	N	N	N	N	N	N	0	4	-4	0	2	3	1	-2
21	-6.0	N	N	N	N	N	N	1	5	-4	0	0	5	6	1
7	-5.8	N	N	N	N	N	N	19	1	18	1	0	5	10	5
51	-5.5	N	N	N	N	N	N	1	1	0	0	1	4	7	3
32	-5.5	N	N	N	N	N	N	3	3	0	0	0	5	7	2
71	-4.7	N	N	N	N	N	N	1	2	-1	0	0	5	3	-2
66	-4.7	N	N	N	N	N	N	2	1	1	0	0	5	6	1
57	-4.5	N	N	N	N	N	N	4	2	2	0	0	4	0	-4
4	-4.0	N	N	N	N	N	N	2	0	2	0	0	3	4	1
75	-3.6	N	N	N	N	N	N	3	1	2	0	0	2	2	0
62	-3.4	P	P	P	N	P	P	5	0	5	0	0	0	3	3
1	-3.3	N	N	P	N	N	N	2	1	1	0	0	2	3	1
35	-3.2	N	P	P	P	N	P	9	0	9	0	0	1	2	1
19	-3.1	N	N	N	N	N	N	10	0	10	0	0	1	2	1
11	-3.0	N	N	N	N	N	N	2	0	2	0	0	1	2	1
43	-2.9	N	N	N	N	N	N	1	1	0	0	0	1	2	1
63	-2.7	P	P	P	N	P	P	19	0	19	0	0	1	6	5
55	-2.2	N	N	N	N	N	N	4	1	3	0	0	1	0	-1
8	-2.1	N	N	N	N	N	N	0	19	-19	0	5	1	0	-1
12	-1.9	N	N	N	N	N	N	7	0	7	0	0	0	2	2
37	-1.6	N	N	N	N	N	N	1	3	-2	0	0	1	1	0
40	-1.2	N	N	N	N	N	N	1	2	-1	0	0	1	1	0
42	-1.0	N	N	N	N	N	N	2	1	1	0	0	1	0	-1
69	-0.6	N	N	N	P	N	N	6	0	6	1	0	0	1	1
59	-0.5	N	N	N	N	N	N	4	2	2	0	0	0	0	0

Exclusive progression: PLR citerion
Progression: non-PLR criterion
Improvement: non-PLR criterion

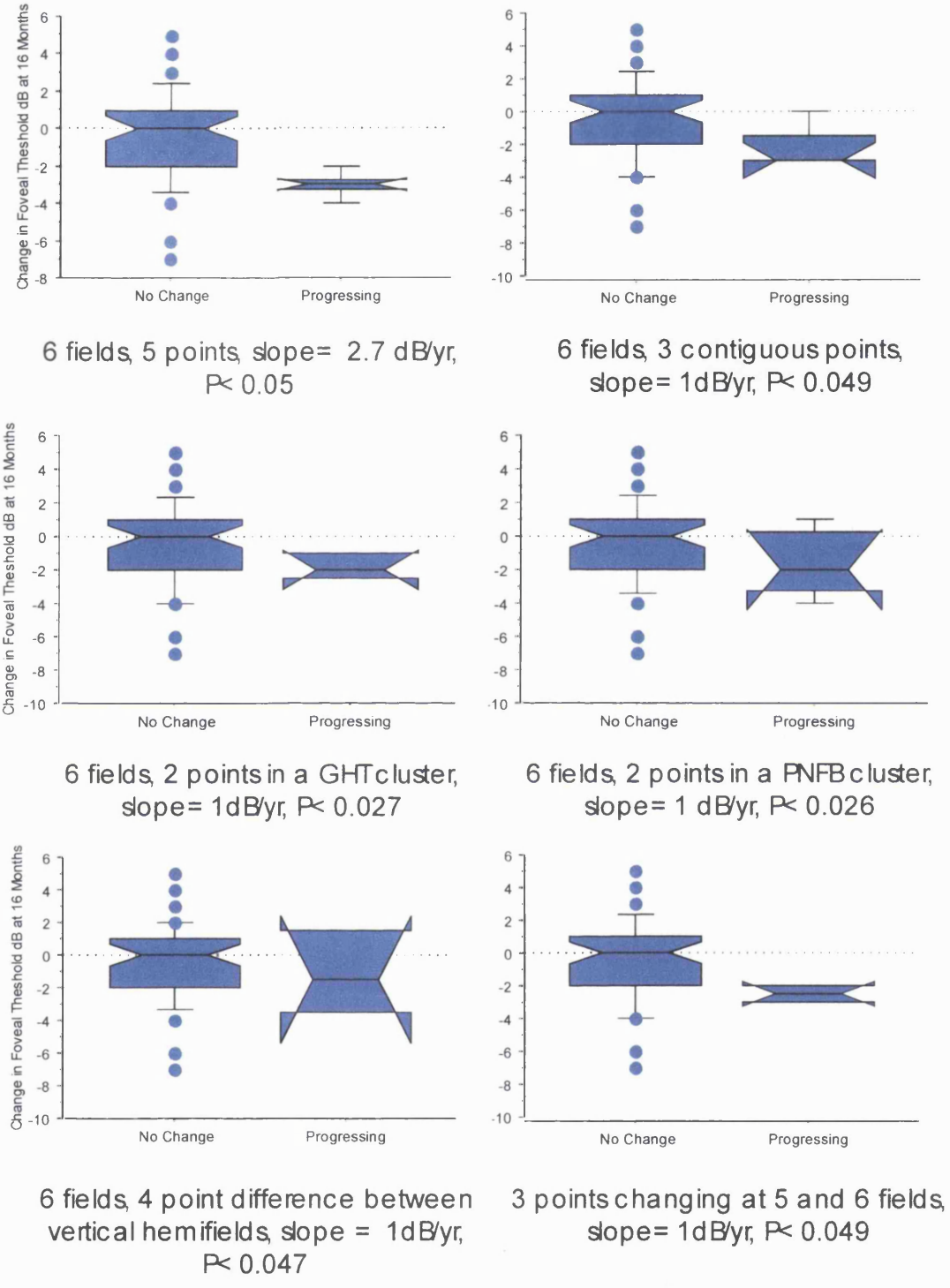
Figure 39 Box-whisker plot explanation



3.6 Progression criteria and Median IOP

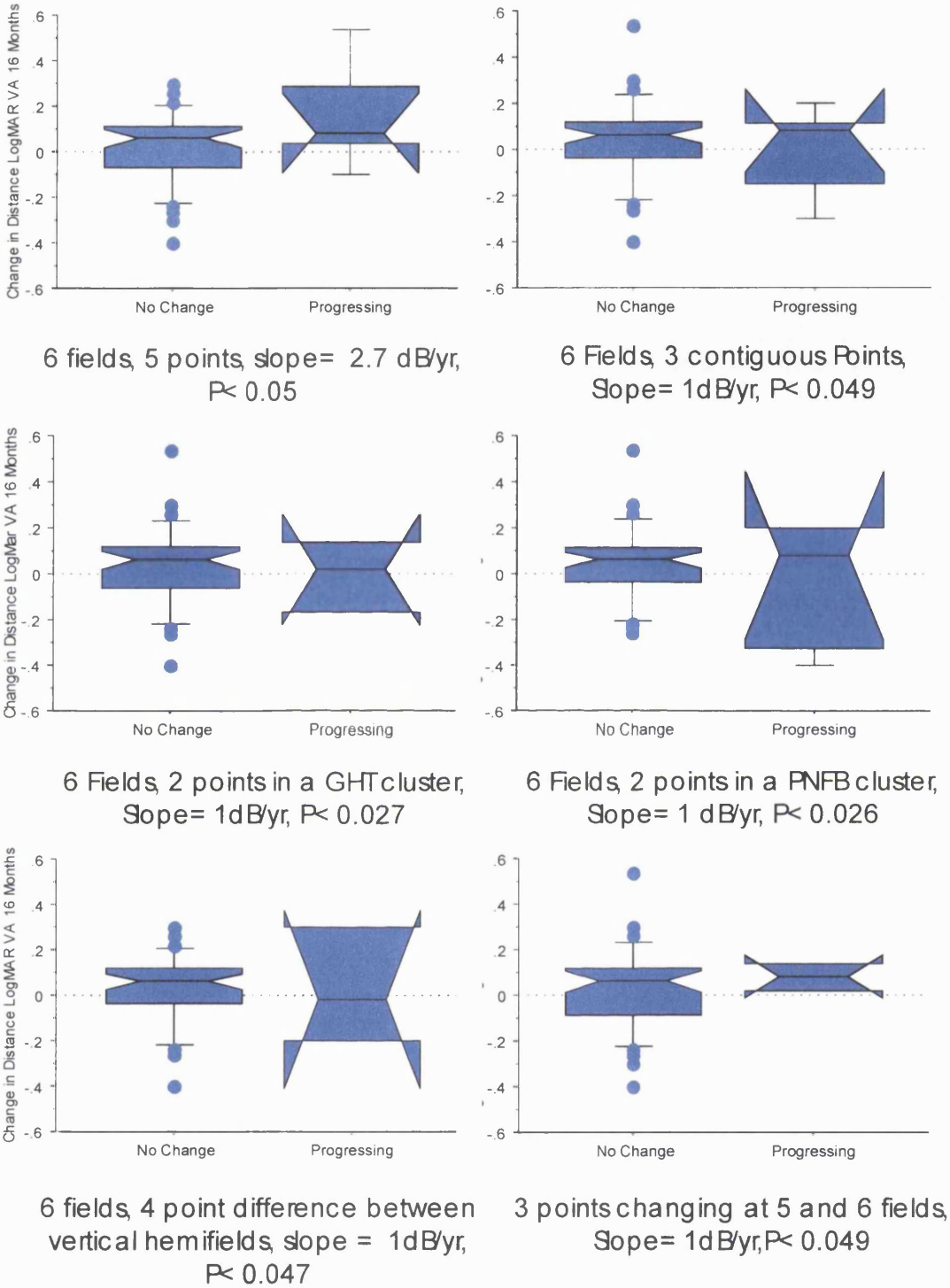
Figure 42 shows the box-whisker plots for the median IOP for progressors and non-progressors as determined by the 6 exclusive progression criteria. There is a noticeable trend that the median IOP is lower in the progressors group than those showing no change. Only with the criterion using a modified perimetric nerve fibre bundle is there no overlap between the 95% confidence intervals for median IOP. The general lower median IOP result is influenced by a single patient with a low IOP in the progressors group. This patient (no. 62) had a mean 16 month IOP of 6 mmHg. With this patient removed from the analysis there is still no overlap between the 95% confidence intervals for median IOP. The eye showed progression using all of the criteria except for the criterion requiring a difference between hemifields of 4 points changing over 6 fields. The same eye also had 5 progressing and 0 improving points using Statpac II analysis. When using percentage fall in IOP instead of mean IOP, a similar picture emerges (figure 43). . There is a general trend for progressors to have had a larger fall in IOP than non-progressors. There is overlap between the 95% confidence intervals for all criteria.

Figure 40 Median change in foveal threshold for the 6 exclusive progression criteria over 6 fields.



Although there appears to be a significant difference in foveal threshold between "Progressing" and "No Change" for some criteria, the small numbers of eyes in the "Progressing" group means that any statistical interpretation must be done with caution.

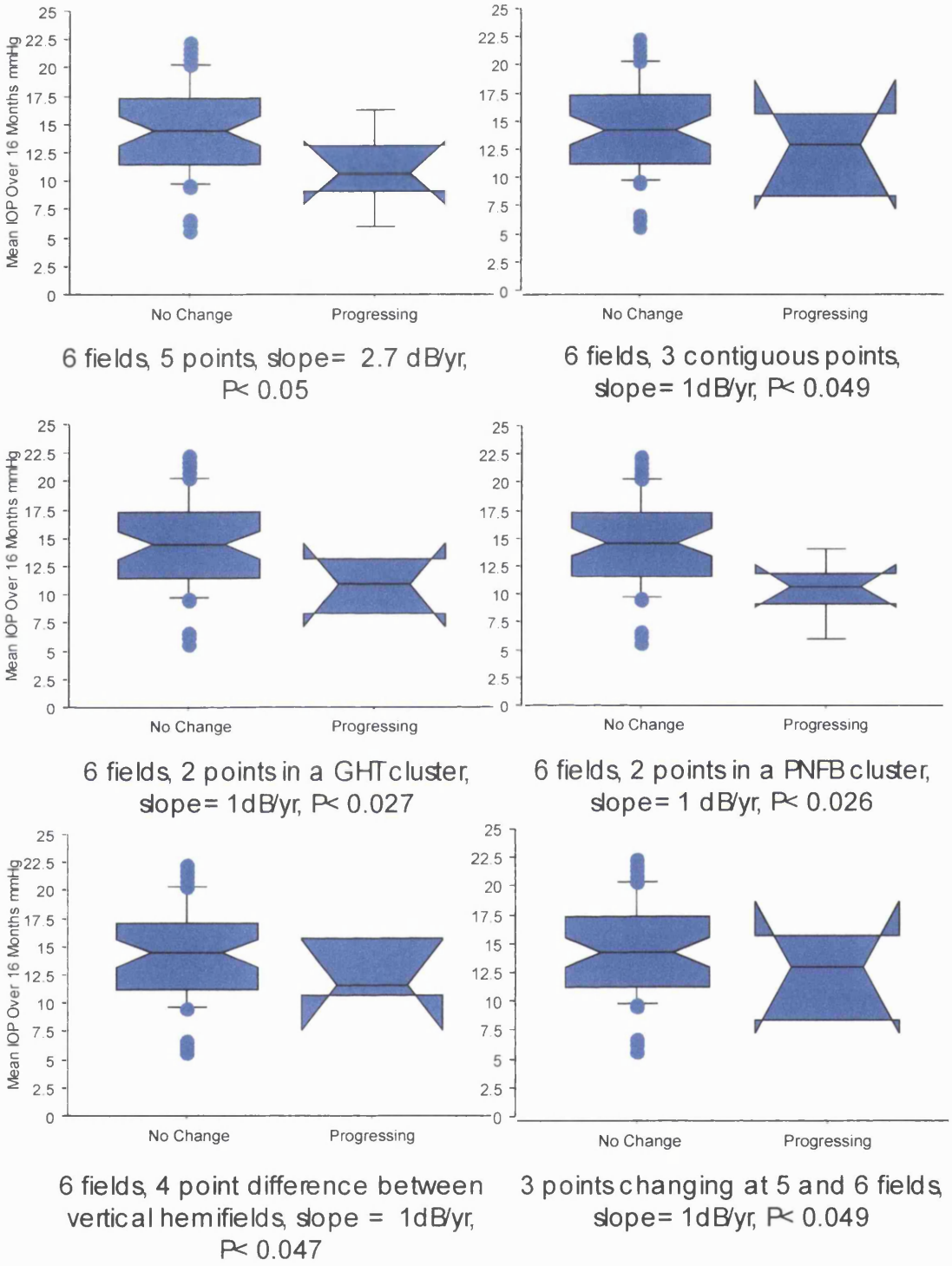
Figure 41 Median change in logMAR VA for the 6 exclusive progression criteria over 6 fields.



No significant differences were detected in median change in logMAR VA for the "Progressing" and "No Change" groups for any criterion. The caveat for fig 34 remains.

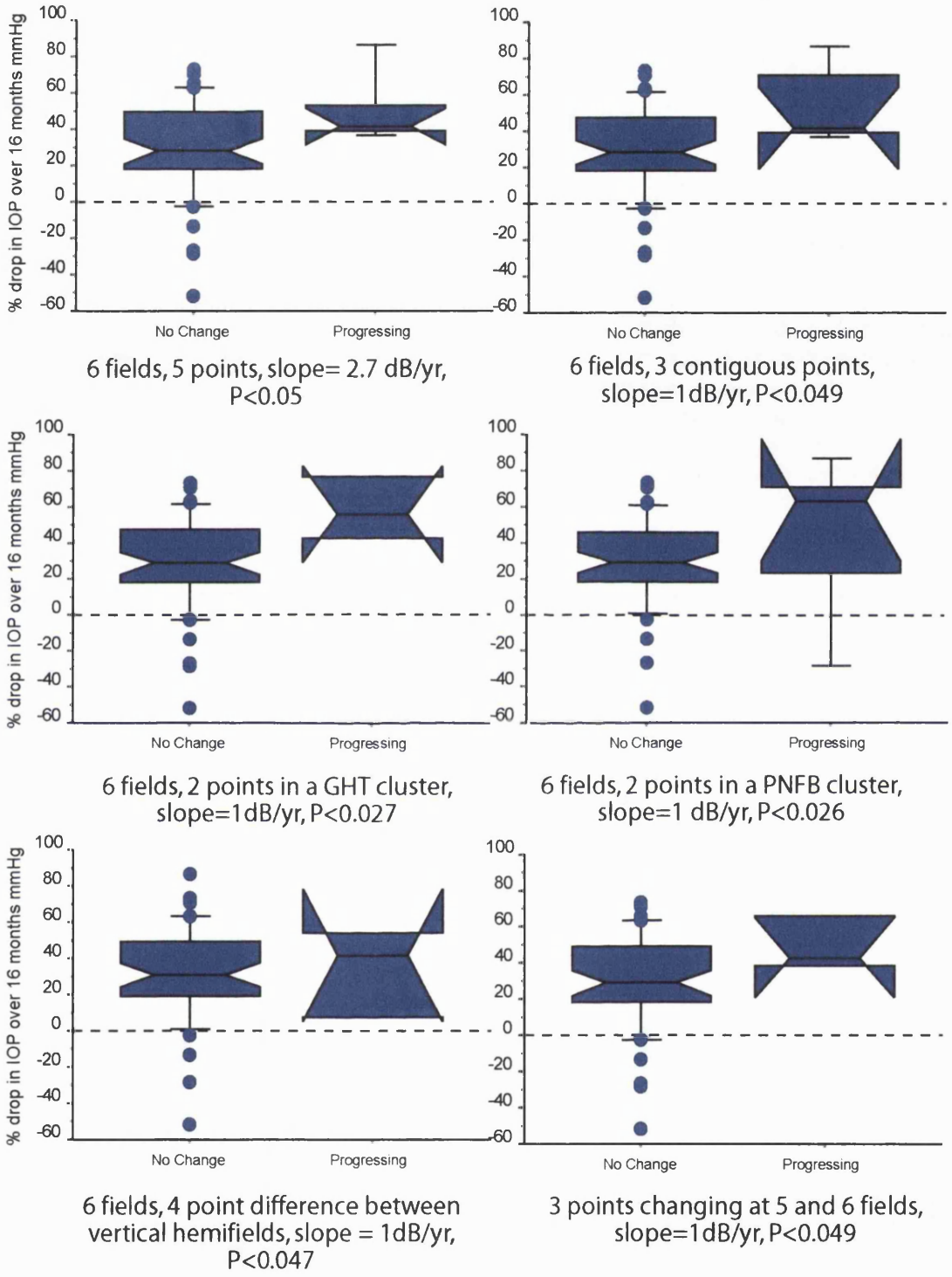
The box-plots reveal no clear pattern. There is for example no consistent detection of greater change in foveal threshold with criteria that do not use spatial relationships between points.

Figure 42 Median IOP for progressors and non-progressors.



6 exclusive progression criteria. The median IOP for the "Progressing" eyes is lower than the "No Change" eyes for all criteria. Only with the PNFB cluster criterion is this difference significant. However the small numbers of "Progressing" eyes limit any statistical conclusions.

Figure 43 Median % fall in IOP for progressors and non-progressors



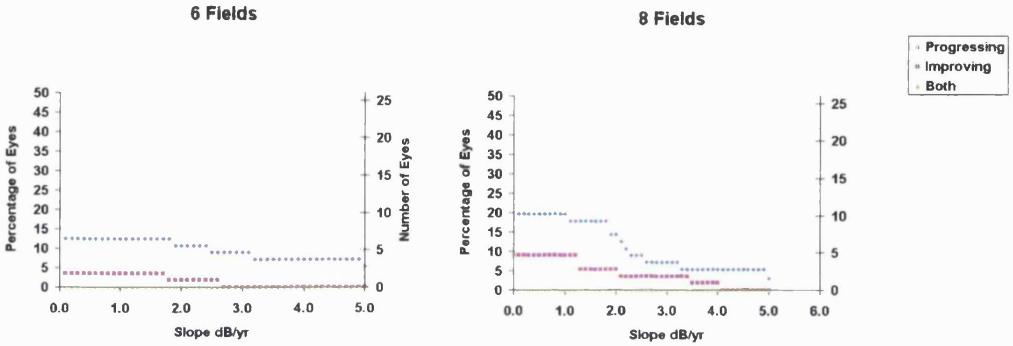
There are no significant differences in % drop of IOP over 16 months for any of the change criteria.

3.6.1 Extended follow up of original 56 patients from 6 to 8 fields

Adding 2 fields to the 6 field analysis produces similar curves with variations in the amount of separation. Exclusive progression is detected using the 5 criteria capable of detecting it over 6 fields.

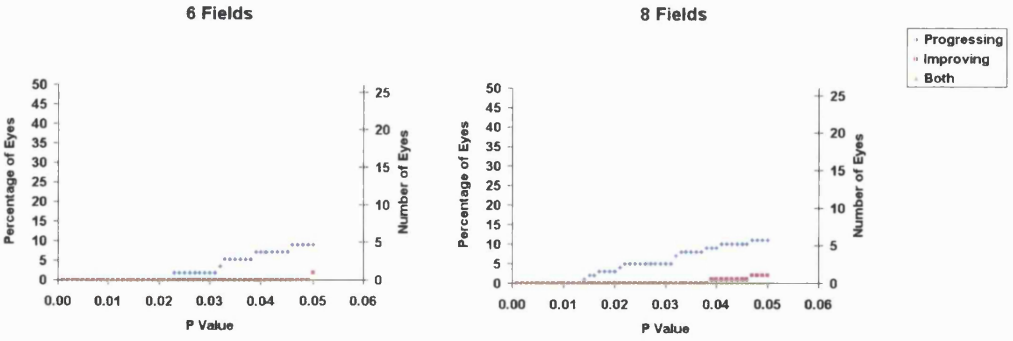
a) 5 points, P<0.05: Less exclusive progression at higher slope values is detected over 8 fields compared to 6

Figure 44 Extended follow up - 5 points changing vs slope, P<0.05



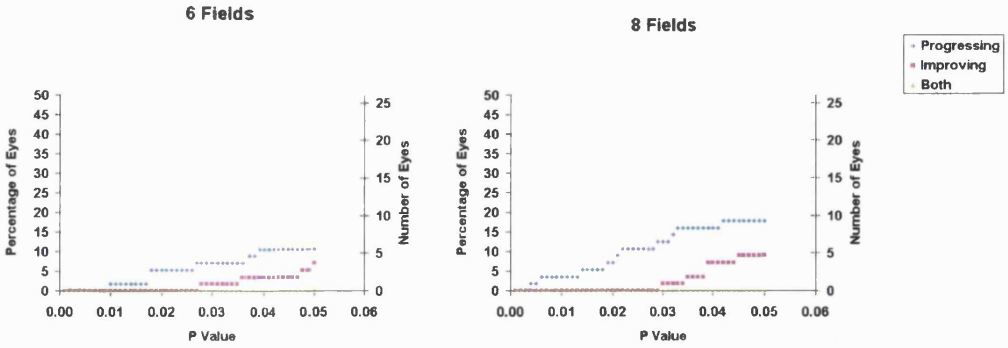
b) 3 contiguous points: The same percentage of exclusive progression is detected at P<0.038 over 8 fields as at P<0.049 over 6 fields

Figure 45 Extended follow up - 3 contiguous points



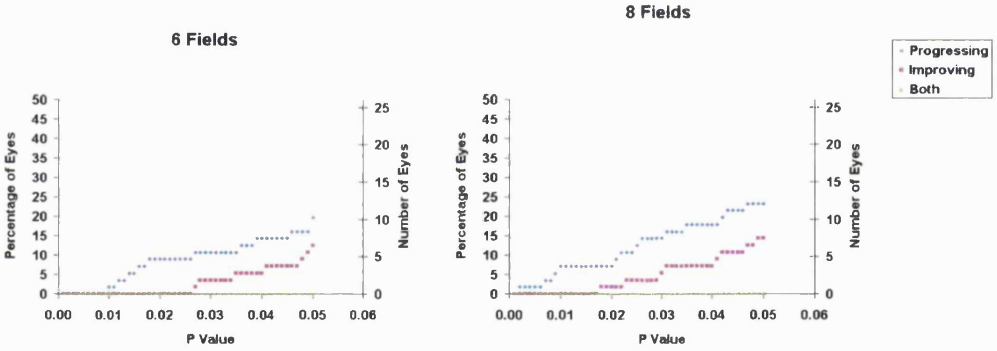
c) 2 points in a GHT cluster: Much more exclusive progression is detected at P<0.029 over 8 fields than at P<0.027 over 6 fields.

Figure 46 Extended follow up - 2 points in a Glaucoma Hemifield Test cluster



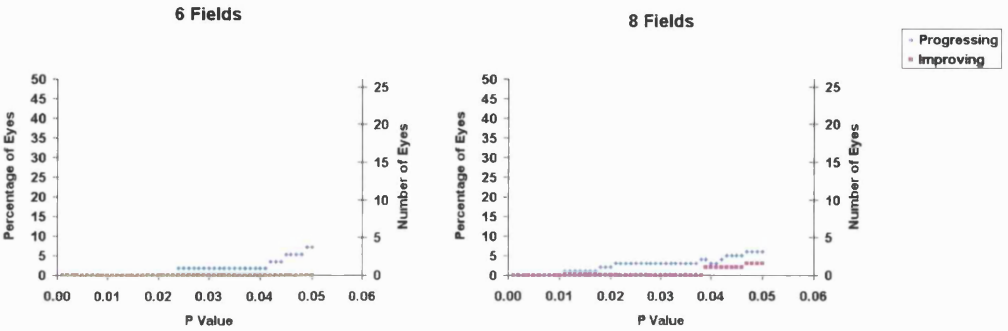
d) 2 points in a PNFB cluster: A reduced amount of exclusive progression is detected at stricter P values over 8 fields at P<0.017 compared to 6 at P<0.026.

Figure 47 Extended follow up - 2 points in a Perimetric Nerve Fibre Bundle Cluster



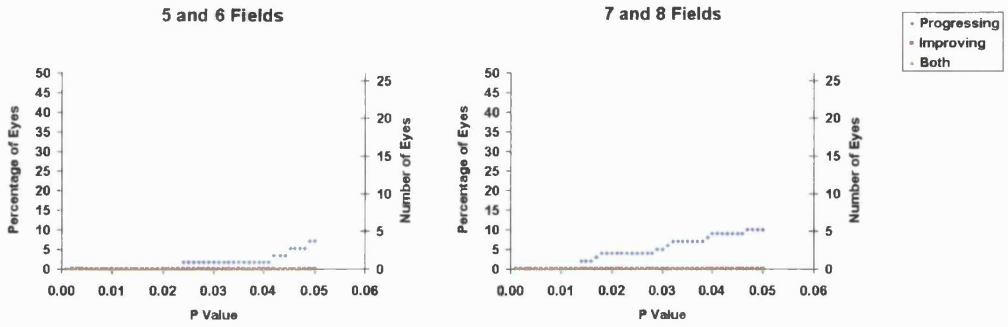
e) 4 point difference between vertical hemifields: Much less exclusive progression is detected at P<0.038 over 8 fields compared to P<0.047 over 6 fields.

Figure 48 Extended follow up - 4 point difference between vertical hemifields



f) 3 points changing at 5 and 6 fields: There is a slight increase in exclusive progression at P<0.05 over 7 and 8 fields compared to 5 and 6 fields.

Figure 49 Extended follow up - 3 points changing over 5 & 6 or 7 & 8 fields



3.6.2 Application of learning set criteria to test patients

The above analyses were performed on the initial 56 patients that entered the trial and completed 16 months of follow up. Subsequently field data from the next 97 patients to enter the trial became available. Baseline characteristics of these 97 are shown in table 14. There were no significant differences in any of the baseline characteristics between the 2 groups of patients.

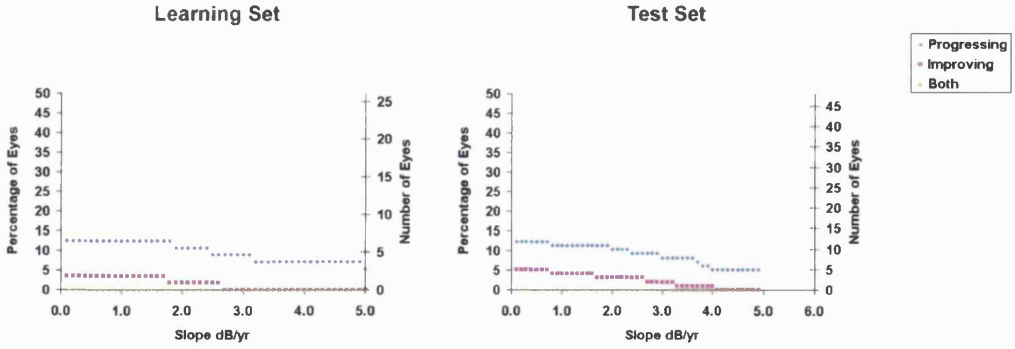
Age	Mean 67 yrs (Range 45-84)
Gender	36F : 61M
Race	Caucasian 80, Black 9, Asian 7
Mean months since glaucoma diagnosed	66.9 (Range 0-348)
Mean Deviation (dB)	-11.6 (Range –0.2 to –29.8)

Table 14 Baseline characteristics of the 97 test patients that formed the test group

In general terms the shape of the curves is maintained when comparing the learning and validation sets, albeit with some differences in their spacing and magnitude.

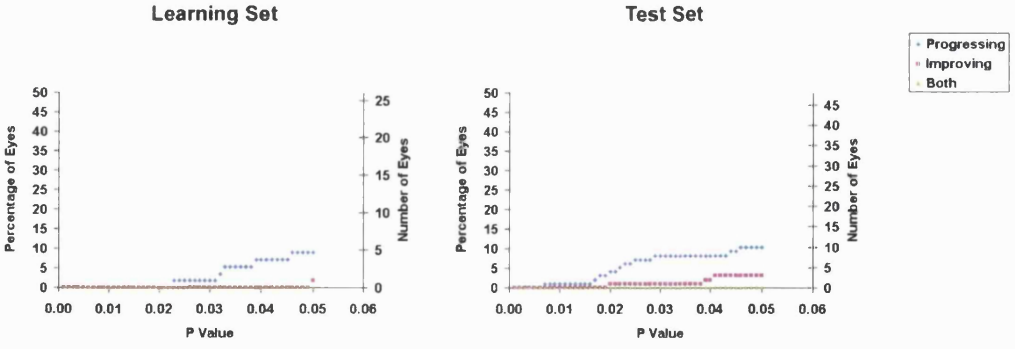
a) Any 5 points P<0.05: the progressing and improving curves are parallel in the learning and test sets. Exclusive progression occurs at a lower critical slope and to a slightly greater degree in the learning set. No detection of "Both" occurs.

Figure 50 Learning/ test sets - 5 points changing vs. slope, P<0.05



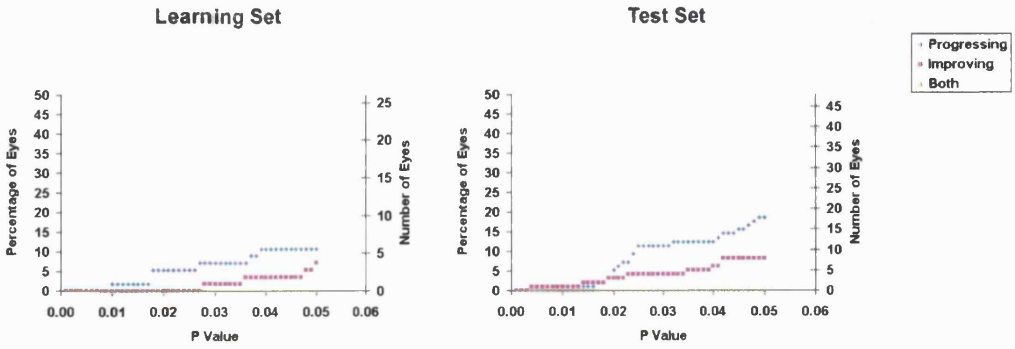
b) 3 contiguous points: More improvement is detected in the test set resulting in only small amounts of exclusive progression being detected at stricter P values than in the learning set.

Figure 51 Learning/ test sets - 3 contiguous points



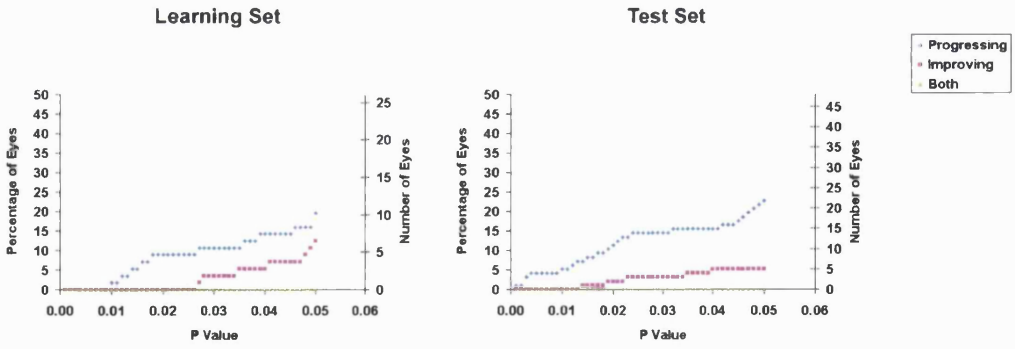
c) 2 points in a GHT cluster: Although more progression is detected at lax P values in the test set, the increase in improvement detection at all P values results in no exclusive progression being detected

Figure 52 Learning/ test sets - 2 points in a Glaucoma Hemifield Test cluster



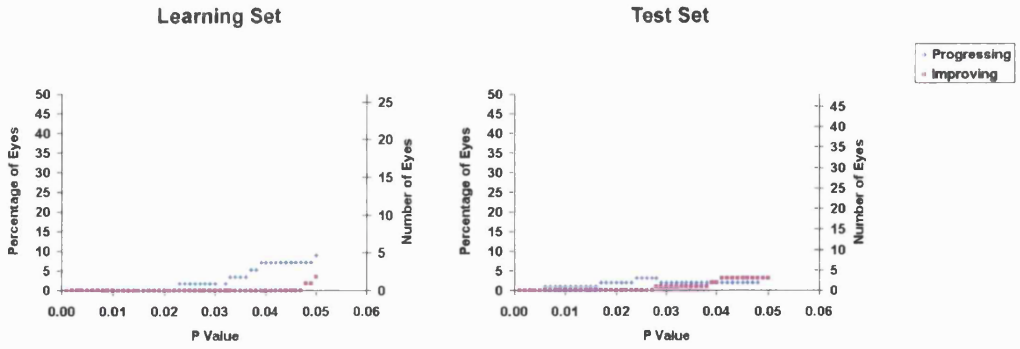
d) 2 points in a perimetric nerve fibre bundle cluster: More progression is detected in the test set with progression being detected down to extremely strict P values P<0.001. Although the improvement curve is flatter in the test set, improvement continues to be detected down to lower levels compared to the learning set. The curves still remain separated with good detection of exclusive progression. The maximum EP is 7.2% in the test set as opposed to 8.9% in the learning set.

Figure 53 Learning/ test sets - 2 points in a Perimetric Nerve Fibre Bundle Cluster



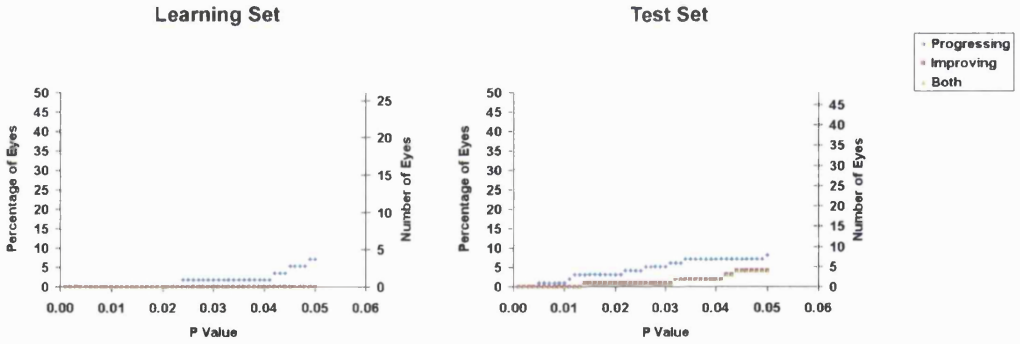
e) 4 point difference between vertical hemifields: Very flat "progressing" and "improving" curves are seen. The fact that the curves are created by subtracting events in one hemifield from the other causes them to cross and uncross. Small amounts of exclusive progression are seen at P<0.027 in the test set, much less than that seen at P<0.047 in the test set.

Figure 54 Learning/ test sets - 4 point difference between vertical hemifield



f) 3 points consecutively changing at 5 and 6 fields: Much more progression and improvement is seen in the test set. This is the only criterion to have seen the re-emergence of the detection of "Both" in the test set. A small amount of exclusive progression was seen at P<0.013.

Figure 55 Learning/ test sets - 3 points changing over 5 & 6 fields



From the analysis of 52 patients followed over 6 and 8 fields and from the analysis 96 patients followed over 6 fields it was possible to select 2 criteria that exclusively detected at least 4 eyes that were exclusively progressing in all 3 data sets. These criteria are:

a) 5 points anywhere in the field, P0.05

b) 2 points in a PNFB cluster, slope = 1 db/yr

Further discussion of the relative merits of these 2 criteria and the implications for using them is contained within the following section.

4 Discussion

4.1 Intraocular pressure reduction following trabeculectomy

To date the MRC 5-FU treatment trial has obtained a significant reduction in mean IOP of 7.3 mmHg at 12 months and 7.2 mmHg at 16 months. No analysis according to treatment arm has been performed because the trial protocol only permits such analysis 1 year after the last patient has been randomised. The general reduction in IOP is slightly less than that seen with other studies on low risk trabeculectomies (Kidd and M 1985; Migdal et al. 1994). Kidd et al found a reduction of 12.1 mmHg at 1 year and Migdal et described a reduction of 21.1 mmHg at 6 months and 20.9 mmHg at 5 years. This may in part be due to methodological differences, Migdal's study required a minimum IOP of 24 mmHg (Migdal et al. 1994), and in part due to which IOP reading was used in the calculations. In this study the pre assessment IOP was used as the baseline IOP rather than the IOP at listing. This was done to minimise regression to the mean. Indeed there was a significant fall in IOP between listing and assessment. It should also be born in mind that if 5-FU works as postulated then it should confer an additional IOP lowering effect to the 50% of randomised eyes that receive it

We found that when looking at the median IOP of progressors and non-progressors for the six exclusive progression criteria only the criterion of 2 points in a PNFB $P<0.026$ produced a median IOP that was significantly lower in those eyes showing progression compared to those that did not. This effect persisted even when the outlier was removed. Should this be interpreted with caution since it does not conform to our expectations? The result may be spurious since one is comparing the median IOP of 4 or 5 eyes with that of 51 or 52 eyes. However hypotony is also known to have a profound effect on macular and optic nerve function (Stamper et al. 1992; Costa et al. 1993). It is possible that hypotony induced cataract, although inspection of the data suggests that in the 3 eyes with a mean IOP of ≤ 6 mmHg no obvious change in cataract occurred. It is possible that eyes that had

hypotony could develop visual field changes giving a "paradoxical" association between progression and IOP. If this effect were to continue it would hamper any overall analysis based on mean or median IOP. It would mean that any influence of IOP on field progression would have to be detected by analysing eyes in IOP bands. The present trial protocol plans for a 3 band analysis, Group 1: < or = 14mmHg, Group 2: 14.01 to 18.99 mmHg, Group 3: > or = 19 mmHg. It is conceivable that any protective effect on visual fields from lowering IOP in group 1 would be masked by progression due to hypotony.

4.2 Changes in media opacity and visual acuity following trabeculectomy

All intraocular surgery is recognised as potentially being cataractogenic. Trabeculectomy is no exception, involving intraocular inflammation, alteration of aqueous flow, changes in the blood eye barrier and prolonged exposure to topical steroids. Thus although our patients were excluded from the trial if they were likely to develop cataract over 3 years it is quite possible that some did go on to develop lenticular opacities. Some studies that have prospectively followed trabeculectomy patients have found low levels of cataract formation in the first 6 months after trabeculectomy. Clarke found no evidence of cataract causing a drop of 2 lines of vision over the first 6 months (Clarke et al. 1990). A larger short term study found that 42/508 (8.3%) eyes dropped 2 or more Snellen lines of acuity in the first 3 post operative months (Costa et al. 1993). Of these only 16 lost vision due to cataract. Other studies, especially those with longer follow up have suggested higher rates of cataract progression: 18.1% of patients developed cataract over 1 year in Robin's study (Robin et al. 1997), and in the AGIS study 37.1% of Caucasian patients randomised initially to trabeculectomy underwent cataract surgery over 7 years of follow up (AGIS 1998). Subsequent analysis of the effect of cataract surgery on the trial outcome felt that the race treatment effect persisted after allowing for the surgery. The analysis was limited by the lack of formal assessment of cataract and by the study design. When assessing this study it should also be borne in mind that the different racial groups had different rates of

cataract formation and had different rates of surgery. Black patients were younger, more likely to be female, have diabetes, have more advanced cataracts, less likely to have surgery, and when they did have it done more likely to have it performed by a resident.

An additional factor to bear in mind is the possible impact of the use of antimetabolites. One study has shown that the use of mitomycin is associated with an increased risk of progression of cataract post operatively (Robin et al. 1997). However Egberts' study of intraoperative 5-FU found no differences in rates of cataract formation between treatment and control arms (Egbert et al. 1993).

In the results presented above there is some evidence of cataract development in that with our modified grading system there was a significant change in nuclear colour (NC) but not nuclear opalescence (NO). Using foveal threshold and logMAR visual acuity as indirect markers of cataract formation is something that should be done with caution since lens opacity is not the only parameter that can effect them. Distance logMAR VA and foveal threshold are only significantly associated with LOCS NO and NC scores at 12 months and not at baseline. Furthermore there are no significant changes in the means of either logMAR distance VA or foveal threshold at 12 or 16 months when compared to baseline.

This paucity of evidence for cataract change may be because, despite expectations, no sizeable change occurred, or because in particular the modified LOCS grading system and sample size prevent subtle changes from being detected. None of the patients underwent cataract extraction during this phase of the study.

Cataract development is clearly an area of great importance for all glaucoma surgery trials. Analysis of the Collaborative Normal-Tension Glaucoma Study Group was complicated by the significantly greater incidence of cataracts (38% vs 14%) in the treatment arm compared to the non-treatment arm (Collaborative Normal-Tension Glaucoma Study Group 1998). Only by censoring data where cataract affected visual acuity could a beneficial effect of IOP lowering treatment on visual field progression be

seen. In the AGIS study (AGIS 1998) rates of cataract surgery are revealed but no analysis is done on their effect on the study's overall conclusions. The main thrust of the study report is that, based on visual function outcomes, there is a racial difference; black patients should be offered the ATT treatment sequence while healthy white patients should be offered TAT (T- trabeculectomy, A – ALT). Table 15 shows the percentage rates for cataract surgery ± trabeculectomy in the study.

Table 15 Numbers of patients requiring cataract surgery in the AGIS study.

	Black		White	
Intervention	ATT	TAT	ATT	TAT
Number randomised	240	211	158	167
Number requiring cataract surgery	85	98	70	76
Percentage requiring surgery	35.4	46.4	44.3	45.5

Large percentages of patients required cataract surgery in all 4 arms of the analysis. The cataract rates were higher in the TAT arm than the ATT arm for blacks, with the same high rate in both arms for whites.

If one assumes that significant cataract formation was occurring then it would be likely that a diffuse change in the visual field would result (Lam et al. 1991; Budenz et al. 1993). One can try to predict the effect of this diffuse change on the pointwise linear regression of fields, one could argue that significant change would be more likely to be detected in "normal" areas where there was lower threshold variability. Thus any true cataract induced change in threshold would not be masked by noise. Diffuse change would furthermore be expected to create artefactual progression with most of the PLR criteria except for the criterion that looks for a difference of 4 points changing between vertical hemifields. Equal amounts of opacity induced change in each hemifield would neutralise each other with this criterion.

Although no statistical analysis was possible for posterior subcapsular cataracts (P). It is clear that a few patients grades for this type of cataract improved following surgery. It seems unlikely that such a cataract would have happened. It may represent a data entry

error, or may reflect the difficulties of grading cataract, especially when some patientshave small pupils preoperatively due to pilocarpine.

4.3 Statpac analysis

4.3.1 Global indices of visual field change

Analysis of global visual field parameters shows no significant mean change over the 16 months of analysis performed. Given the slow rate of change of similar parameters noted in other studies it would be more remarkable if significant change had been noted. In the Glaucoma Laser Trial the mean change in decibels per test location improved for the first three years of the trial (Glaucoma Laser Trial Research Group 1995). It has also been shown that global parameters improve following trabeculectomy (Gandolfi 1995).

Table 13 shows the correlation between change in IOP and Statpac global indices over 16 months. Although the figure for MD has a significance of 0.034 at 16 months this should be interpreted with caution. Given the 8 analyses performed it would be prudent to apply a Bonferroni correction thereby reducing the level of significance to 0.00625 (Altman 1991). This would mean that none of the changes in global indices showed a significant correlation with change in IOP. This is not surprising given the lack of significant change in the individual Statpac variables themselves. As discussed in section 1.2 the evidence on whether lowering IOP has an effect on visual field progression is not conclusive.

Linear regression of global Statpac indices showed no significant change for MD, PSD, SF, and CPSD. This matches the finding of others who have either not been able to detected significant change using this technique (Chauhan et al. 1990) or who found that low amounts of change were detected with it (Nouri-Mahdavi et al. 1997). In an analysis of 191 patients Smith et al were able to show that only 12.6% showed significant MD slope and 14.1% showed a significant slope in CPSD (Smith et al. 1996).

4.3.2 Pointwise analysis using Glaucoma Change Probability analysis

Analysing change over 6 fields (figure 18) shows that a large amount of progression and improvement are detected. Glaucoma Change Probability analysis merely involves a pointwise comparison of the last field in a series with the mean of the first two. If one were to arbitrarily select a progression criterion for this field series it would need to be borne in mind that even a strict one requiring large numbers of points to change would detect just as many patients improving as progressing. Using the mean of the thresholds from the initial 2 fields does help to reduce inter test variability, although Heijl's early work used the mean of the first and last pair of fields (Heijl et al. 1989). Anticipating this variability subsequently, Heijl et al were of the opinion that "since the maps use the 5% and 95% limits, random variability should be expected to give a few significant points even in stable eyes" (Heijl et al. 1990). What is still not clear is exactly what are the normal limits for significant points.

If consecutive change is required over the final three fields then interestingly more patients are shown to be improving than progressing (figure 38). This is in contrast to the AGIS scoring that suggests that no change is occurring (Section 4.3.3). A previous study comparing pointwise linear regression and Glaucoma Change Probability analysis only used a change criterion of the same point changing in any 2 of the last 3 fields. With this criterion comparable sensitivities and specificities in detection of glaucomatous change compared to PLR were achieved (McNaught et al. 1996).

4.3.3 AGIS scoring

Out of 56 patients followed for 16 months 5 showed progression and 2 showed improvement. The definition of change was an increase or decrease in score of 4 or more. This figure was derived from the AGIS group who state that retesting of fields produces yields a change of 4 less than 5% of the time (AGIS 1994). Thus with random change one

might expect 2-3 patients to show improvement and 2-3 to show progression. So only 2 or 3 of the patients are progressing, though it is impossible to say which of the 5 truly are.

These figures for significant change in AGIS score, especially improvement, are markedly different from those seen in a previous study using AGIS grading where 7.5% showed progression and 11.9% showed improvement (Katz 1999). It may well be that the patients' level of field experience in the MRC 5-FU trial was greater and that this contributed to the reduced levels of improvement.

The change in mean AGIS score from 7.86 to 8.21 over 16 months was not significant. In the AGIS study the change in score in Caucasian patients randomised to TAT was approximately +0.15 over 18 months and was not significant (AGIS 1998). This probably reflects the inability of the scoring system to detect small change.

In re-evaluating the CNTG progression criteria Schulzer et al produced an estimate for the true rate of field progression of 1.31% per annum using their criteria (Schulzer et al. 1991; Schulzer 1994). This difference in detection rate is due to a difference in patient characteristics, and the effect that different scoring systems are known to have on progression detection (Katz 1999).

4.4 PLR progression criteria

Crucial to the outcome of the MRC 5-FU study will be the analysis that shows whether patients exposed to intra operative 5-FU were less likely to show loss of visual field than those patients who received a control solution. Indeed the power calculation of the whole study is based on assumptions about rates of field change 3 years after surgery (Section 4.5.3). This outcome measure in turn relies on how one sets about defining visual field progression. This thesis has set about examining automated visual field data in an attempt to generate meaningful progression criteria. The current trial protocol defines progression as 1 point changing with a significant slope of 1 dB/yr with $P < 0.05$. This may turn out to be an appropriate criterion when data for the patients with 3 years of follow up

is analysed. However from the data shown here it is clear that such a criterion is probably over detecting change over shorter periods of follow up (figure 19).

Applying stricter criteria predictably reduces the overall detection of change. What is less predictable is the effect of different criteria on change detection. From the initial analyses it can be seen that manipulating the critical slope in the range 0-5 dB/yr is a rather insensitive way of influencing change detection. In contrast manipulating the P value in the range 0.001 to 0.05 has a more profound effect. Obviously one is not comparing like with like, however these ranges in which the parameters were manipulated represent the commonly used ranges. In assessing points with significant slopes Wild calculated the percentages of points with significant slopes that were 2.5, 5 or 10 times the magnitude of the normal decline in sensitivity with age (Wild et al. 1997). It could be argued that even higher slope values might have a greater differential effect on the detection of progression and improvement. However assuming a "normal" age related decline insensitivity of 0.1 dB/yr then 5 dB/yr already represents 50 times this "normal" rate. It would be useful to be able to detect a 1dB/year rate of decline. Such a rate would take a point from full (30 dB) to no sensitivity in 30 years, easily within the lifetime of a glaucoma patient. It could be that with longer follow up slope criteria in the 0-5 dB range will be more discriminating in detecting change.

Previous work has noted the detection of improvement using PLR (Smith et al. 1996; Bhandari et al. 1997; Katz et al. 1997; Wild et al. 1997), however some early work assumed that those patients without points with significant negative slope should be labelled as not changing (Noureddin et al. 1991; Birch et al. 1995). What is clear from the data in this study is the appearance of patients showing both progression and improvement when lax criteria are used. I had initially thought that simply by tightening change criteria to remove the detection of "Both" then appropriate change criteria could be defined. However it became clear that, with some additional modification, criteria that exclusively detected progression could be created.

Either by requiring more points to have a significant slope or by requiring a spatial relationship between the points a reasonable number of patients can be detected as exclusively progressing. From the analysis of the initial 56 patients it is not possible to tell if one of the criteria displaying exclusive progression is obviously preferable to the others. From figure 38 one can see that criteria capable of detecting exclusive progression do so on different patients with some overlap in patients between criteria.

Equally one can see that using spatial filtering in an attempt to deal with point variability did not help exclusively select progressing patients. Previous work with Gaussian spatial filters has been shown to improve the accuracy of predicting progression using PLR (Crabb et al. 1997). No work was done on looking at the ability of spatial processing to separate stable and progressing glaucoma subjects. It is felt that more accurate progression of future progression reflects a more valid assessment of change within existing data (Crabb et al. 1997). In this study spatial filters were not helpful. It may be that the effect of the filters was too small to be detected and that alternative filters would be more helpful. Further work might look into filters that take into account either the cluster arrangements used above or the layout of the nerve fibre layer.

In the Normal-tension Glaucoma Study Group report the original Statpac II criteria were found to over diagnose progression (Schulzer et al. 1991; Schulzer 1994). It was only by repeating field tests to confirm progression that false positives were reduced to an acceptable level. In this analysis requiring fields to show consecutive change over 5 and 6 field analysis has not been shown to be helpful. It may be that requiring consecutive change over 5, 6, 7, 8, or more fields might produce acceptable levels of exclusive progression.

Removal of 2 (McNaught et al. 1996; Katz et al. 1997) or 4 (Chauhan et al. 1990) points around the blind spot prior to PLR has been used by some groups but not by all (Nouri-Mahdavi et al. 1997). The reason given for the removal of the points given is the increase in variability in points adjacent to the blind spot (McNaught et al. 1996). Again

blind spot removal was not helpful using this dataset. Earlier workers have dealt with the possibility that edge points a) may be lens rim artefacts or b) show increased variability by either removing them from the analysis (Birch et al. 1995) or by requiring stricter progression criteria (McNaught et al. 1996). Stricter progression criteria for edge points had little effect on the overall detection of change in this study.

All the criteria that did not compare the upper hemifield with the lower one could be open to the criticism that they were merely detecting progression due to cataract. Increasing media opacity will reduce the sensitivity of all points. Points that are spatially related will continue to be thus depressed. It was interesting to see that the 4 patients detected as exclusively progressing using a criterion of 2 points changing in a GHT cluster are also found to be progressing if the criterion is a difference of 2 significant points between symmetrical GHT clusters. This is an effect one would not expect if one were observing purely a cataract related effect. Although other work has assessed the sensitivity and specificity of cross-meridional analysis in separating glaucomatous fields from normals (Katz et al. 1991), no study has previously looked at such a technique in assessing glaucomatous progression.

The comparison of progression status using PLR and Statpac II in figure 38 shows that there was some degree of overlap in detecting exclusive progression between the 2 techniques. It is not really fair to compare the techniques too much; no spatial criteria were applied to the Statpac analyses. However it is fair to say that the glaucoma probability change analysis in Statpac II detects a large number of individual points showing progression and improvement (table 1).

The random analyses were performed to try and ascertain the levels of change detected that were due to noise. If one compares figure 19 and figure 33 which display the measured amounts of change and the change seen in randomised fields then the improvement curves are very similar regardless how many points are required to change. With the progression curves the amount of measured progression is greater than the

amount of randomised progression when more than 1 point is used as a change criterion. This suggests that some of the progression detected using even a non-spatially related change criterion is real. Equally the similarity between the improvement curves in figure 19 and figure 33 highlights how much change can be due to noise.

If one looks at the mean slope of significantly changing points in the non-randomised and randomised series (figure 34 and figure 35), the measured slope in the non-randomised series tends to that of the randomised one as the P value approaches $P<0.05$. The same is true when looking at the mean baseline sensitivity of changing points in the randomised and non-randomised series (figure 36, figure 37). This suggests that change seen over the field as a whole at the $P<0.05$ level is likely to be due to noise. What is also seen from Figure 37 is the effect that progression is only seen in points with a baseline sensitivity of 26 dB because if their baseline sensitivity is lower there is insufficient "room" for their sensitivity to fall further. This is sometimes known as the "false floor" effect.

4.5 Application of criteria to additional data

4.5.1 Extended follow up of original 56 patients from 6 to 8 fields

It is not surprising that very similar patterns of change are detected when 2 additional fields are added to the initial 6. Seventy five percent of the data in an 8 field analysis is used in the 6 field analysis.

Over 8 fields the progression and improvement curves are shifted leftward for all criteria (or rightwards when plotted against slope). For any given P value the effect on change detected over 8 compared to 6 eyes is more complicated since the slope of the change curves is also affected. Progression detected over 6 fields due to random change would be expected to decrease over 8 fields when analysed using PLR. Thus real changes in retinal sensitivity are probably being detected.

The change in the improvement curve relative to the progression curve is not consistent, i.e. the spacing between the curves varies for each criterion. This lack of obvious trend makes it difficult to draw any inferences. What is still clear is that random noise and genuine learning still leads to the detection of significant levels of learning when using spatial criteria with strict P values to analyse 8 fields performed over 2 years. What is also clear is that 2 criteria appear less attractive over 8 fields:

a) 4 point difference between vertical hemifields, slope = 1 dB/yr

b) 3 points changing over final 2 fields, slope = 1 dB/yr

4.5.2 Application of learning set criteria to test patients

An obvious criticism in deriving criteria by using a single data set is that one is likely to be able to derive something that appears to meet your requirements. More interesting is what happens when these same criteria are applied to a fresh data set. Moving from the learning to test sets again one sees a general leftward shift in change curves (again rightward when change is plotted against slope), with again a variable change in the spacing between progression and improvement curves. Four criteria perform badly on the test set, detecting no or minimal exclusive progression:

a) 3 contiguous points, slope = 1 dB/yr

b) 2 points in a GHT cluster, slope = 1 dB/yr

c) 4 point difference between vertical hemifields, slope = 1 dB/yr

d) 3 points changing over final 2 fields, slope = 1 dB/yr

4.5.3 Optimal progression criteria

Choosing a sensible progression criterion is now easier after the analysis involving additional data. By using extended follow up and by comparing the learning and test sets one is left with 2 potential criteria:

a) 5 points anywhere in the field, $P<0.05$

b) 2 points in a PNFB cluster, slope = 1 dB/yr

With both of them either a critical slope or a P value is expressed. Maximal exclusive progression is detected by varying the P value or slope respectively. Of the 2 the latter would be the more attractive to choose. Both detect exclusive progression at a rate of 8.93% over 6 fields, however the criterion "2 points in a PNFB" is less negatively affected than "5 points anywhere in the field, $P<0.05$" when the criteria are applied over 8 fields or to the test set (table 16).

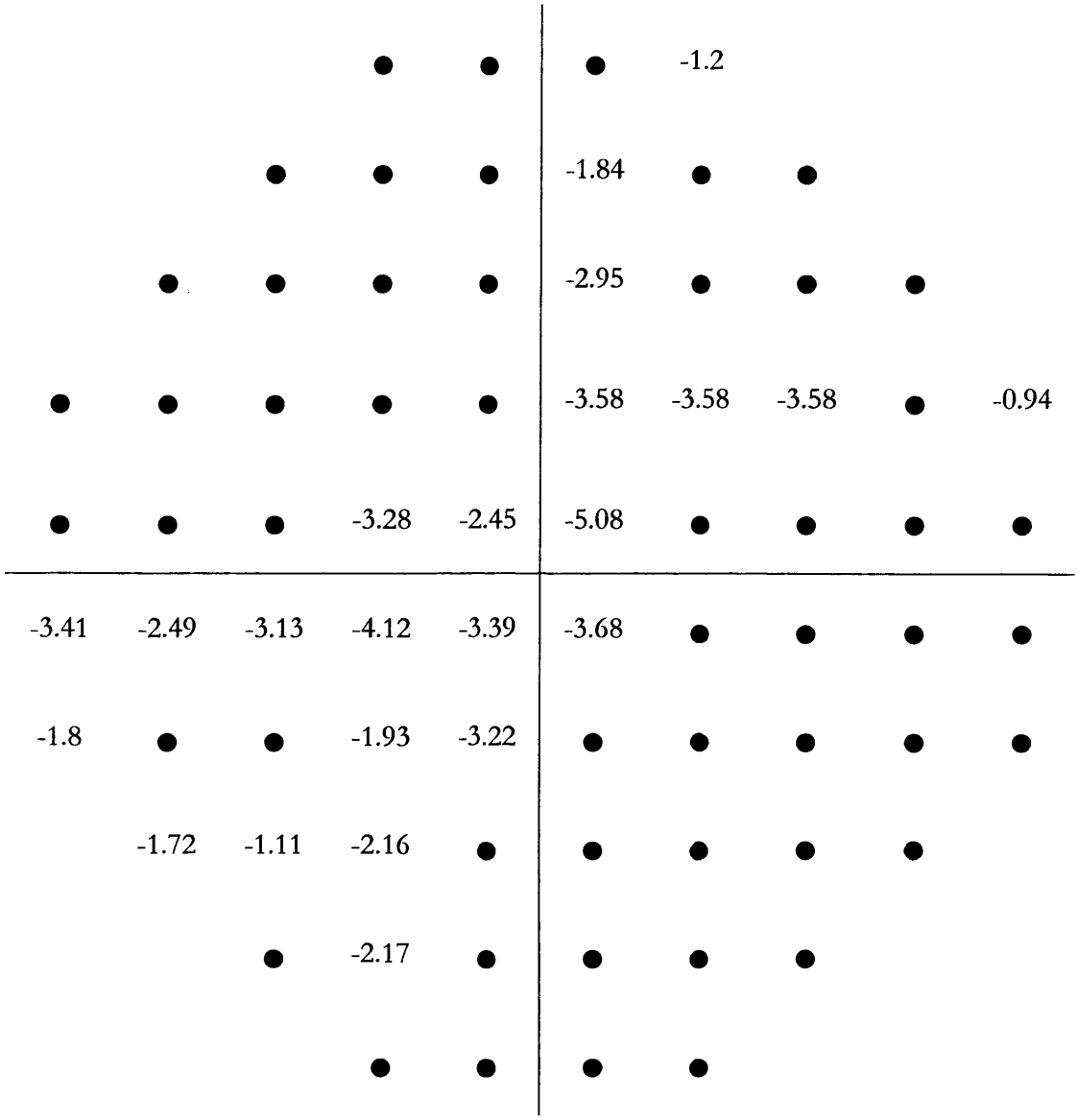
Table 16 Percentage of exclusive progression detected using 2 criteria, and 3 data sets.

	6 fields/ Learning set	8 Fields	Test set
5 points anywhere in field, $P<0.05$	8.93%	5.35%	5.15%
2 points in a PNFB	8.93%	7.14%	7.20%

Furthermore the criterion "5 points anywhere in the fields, $P<0.05$" requires the critical slope to be high, it requires that the critical slope to be increased from 2.7 dB/yr to 4.1 dB/yr to obtain maximal exclusive progression with extended follow up (figure 34); and the number of points involved would mitigate against detection of progression in an eye with advanced loss. Although previous workers have found points with similar slopes, their frequency is low.

Figure 56 shows the mean loss of threshold sensitivity in decibels per year in those points found to have a significant slope $P<0.001$, only 2 points have a slope >4.1 dB/yr. Even though this represents a different P value from the one used here we can see from figure 34 that very little change in the mean slope of changing points occurs as the P value is moved from $P<0.05$ to $P<0.001$.

Figure 56 The average loss of threshold sensitivity in dB/yr.



The slopes of individual locations where a significant loss occurred range from –0.94 db/yr to –5.08 db/yr (Katz et al. 1997).

Finally we know that the mean slope and baseline sensitivity of changing points tend towards the values obtained in randomised fields as the P value tends towards $P<0.05$. This is the level of significance used by the criterion "5 points changing anywhere in the field". Thus the criterion "2 points in a PNFB" would be appear to be a better choice of criterion.

It is not possible to validate new criteria for visual field progression against a gold standard. Even the optic disc data that will be generated by the trial will provide an interesting comparison but will not be the absolute arbiter of glaucomatous progression.

By exploring the data in the ways set out above it will be possible to state that by using a certain criterion with a specified P value a given proportion of patients showing significant change will show progression. One could foresee a criterion that at worst had a 5% improvement detection rate when applied at a specified time point. Criteria could be developed for a trial without looking at the treatment code prior to analysing the data from separate treatment arms. To avoid bias one would probably analyse a random subset of the data and then apply the criteria to the whole dataset. Ultimately this is a technique for developing criteria for large groups of patients. Their applicability in the clinical setting to an individual patient will always be clouded by the clinician's knowledge of the extent of the pre-existing visual field defect, and other parameters particularly IOP and media opacity.

Previous techniques to measure field progression have utilised knowledge of long term fluctuation to establish limits of normality. Statpac II relies on retesting subjects with glaucoma (Heijl et al. 1990) to produce levels of probability that a change in a point is real. Similarly as mentioned above (AGIS 1994) the AGIS grading system relies on the fact that retesting produces a change in score of 4 on less than 5% of occasions. Although one would like to evaluate a grading system prior to introducing it into a study, a system has also been re-evaluated mid trial (Schulzer et al. 1991; Schulzer 1994). In the Collaborative Normal Tension Glaucoma study investigators became concerned that visual field end points were being reached too frequently. Their endpoint had been developed to detect the smallest possible change, in developing it they had been concerned about the low specificity of the end point. To reduce it they labelled any change detected with a single field as tentative and then required a further 1 or (if it did not also show progression) 2 fields to confirm the progression. By only repeating fields if progression was tentatively

seen it was calculated that specificity would be improved while keeping sensitivity high. However when the trial started running it was found that the specificity of an individual test was unexpectedly low at 84.3% and that the overall specificity from confirmatory testing was 95.5%. Using a mathematical model Schulzer et al were able to calculate the effects on detection rate of additional testing. If progression was again detected 3 months after the initial detection and if a further 1-2 fields again confirmed progression then the patient was said to have "definitively documented progression". Using 2 blocks of 2-3 fields 3 months apart the sensitivity was 100% and the specificity 98.3%. Furthermore by analysing 2 groups of data, those not yet randomised in the trial and those randomised to control it was possible to use the mathematical model to generate an estimate of the rate of progression of 1.31% for the non-randomised and 0.51% for the randomised patients.

In analysing the data from this study we have found that the existing progression criteria are too lax. Table 17 shows the estimated rate of field progression calculated as

Table 17 Power calculation from the MRC 5-FU Filtration Surgery Study

	Estimated % of patients with fields worse		Total required in both arms	Comment
	Control	5-FU		
Year 1	10%	5%	1161	Not practical
Year 2	17%	8%	563	Not practical
Year 3	25%	11%	312	Practical

The table, taken from the trial protocol, shows the detectable differences in rates of field progression, the minimum number of patients required to detect such a difference and whether the planned recruitment of 367 patients would leave sufficient follow up to detect the difference at annual time points. The figures are estimates based on field data from a pilot 5-FU study (Lanigan et al. 1994), a previous trabeculectomy study at Moorfields (Hitchings et al. 1994; Migdal et al. 1994), and data from other studies (Kolker 1977; Werner et al. 1977; Greve and Dake 1979; Rollins and Drance 1981). Based on previous studies the protocol assumes an annual drop out rate of 5% which means that to have followed up 312 patients over 3 years a total of 367 would have to be recruited. (Migdal et al. 1994; FFSG 1996).

part of the trial protocol. The figures assume that 2/3 of progression will occur in the control arm with the rest in the 5-FU arm. The figure of 312 patients required at 3 years to detect a significant difference in field loss between the 2 arms was used to define the minimum number of patients required for the trial. Assuming a drop out rate of 15% over 3 years, 367 patients would need to be recruited to leave 312 patients being followed at 3 years.

Table 18 reviews the rates of progression available from the original 56 patients. As with the initial power calculation, it is assumed that 2/3 of the progression occurs in the control arm, with 1/3 occurring in the 5-FU arm. It uses a power calculation (Fleiss 1981) to assess the effect of adopting either the PLR or AGIS criterion.

Table 18 The effect of applying measured rates of progression with new progression criteria to the 5-FU trial's power calculations

Analysis technique	Change criterion	Duration of follow up	Measured rate of change	Total required in both arms
PLR	2 points in a PNFB, P<0.026	16/12	8.93%	961
AGIS	4 points or more	16/12	8.93%	961
PLR	2 points in a PNFB, P<0.017	24/12	7.14%	1230

The total required refers to the number of patients that would need to be analysed in order to detect a significant difference between the 5-FU and control arms assuming that 2/3 of progression occurred in the control arm and 1/3 in the 5-FU arm.

Over 16 months using an optimal PLR criterion ("2 points in a PNFB"), or AGIS scoring, 961 patients would be required. The original trial power calculation produces a figure for 12 months of 1161 patients. Neither figure is achievable given the current trial protocol. Using the same PLR criterion over 24 months 1230 patients would be required. It should be noted that, with the data in this thesis, using the optimal PLR criterion "2

points in a PNFB", overall change decreased with increased follow up. This is in stark contrast to the original estimates that calculated the overall rate of progression would increase from 15% to 25% from years 1 to 2. So although the original PLR criterion over detects change, the optimal PLR criterion defined here would require increased numbers of patients in the trial to detect a significant change in visual fields between the 2 treatment arms.

5 Conclusion

In this thesis I have shown that existing criteria for detecting change in automated visual fields using pointwise linear regression can be improved for my data set. The amount of change detected is so large as to make meaningful interpretation of a trial's results impossible. I have explored the effect that varying pointwise linear regression parameters have on the detection of change. In modifying the PLR change criteria, particularly by spatially relating points, it is possible to produce criteria that exclusively detect progression.

By extending the analysis of the initial or "learning" data set to the analysis of 1) additional fields and 2) a separate data set it has been possible to narrow down the potential choice of progression criteria. Only 2 criteria are robust enough to be applicable over extended follow up and to a new data set. Of these 2 it is the criterion involving 2 points changing in a perimetric cluster that appears to offer the best option for exclusively detecting progression. Selecting a single optimal P value for this or any other criterion is not possible. One needs to analyse the whole data, or a randomly selected subset, at each time point and derive the P value that maximally generates exclusive progression. This P value may well need to be changed as the length of follow up is altered. However if the technique used to do this is clearly defined then meaningful criteria can be derived.

6 References

7 Abstracts presented from this thesis

ARVO 1998, Fort Lauderdale, USA - poster

The effect of test number and progression criteria on the detection of field changes in trabeculectomy patients

M.R.Wilkins, T.Lowe, A. Kotecha, P.T.Khaw, F.W.Fitzke

Wound Healing Research Unit, Department of Pathology, Institute of Ophthalmology; Glaucoma Unit, Moorfields Eye Hospital, London, United Kingdom

PURPOSE. To study the effect of changing field number and progression criteria on the numbers of patients showing worsening or improving visual fields following trabeculectomy. **METHOD**. 47 patients listed for trabeculectomy had at least 2 Humphrey Mk II 24-2 fields prior to surgery followed by further tests at 3 month intervals in the first post-operative year. Field analysis was performed using pointwise linear regression. Field series were analysed using the 4 post operative tests with and without the second of the preoperative tests. The effect of using a simple 3X3 Gaussian filter to blur individual fields was compared with non blurred fields. The presence of 1 or 2 changing points were used to label change in a field series. Significant point change was defined as a slope >1dB/year $P<0.01$ or $P<0.05$. **RESULTS**. Where a single point was used as the condition for change then addition of a 5th field had no effect on the numbers of patients progressing or improving in a series. If a 2 point change criteria was used then adding a 5th field significantly increased ($P<0.05$) the number of fields showing deterioration while leaving the number of improving fields unchanged. The use of a Gaussian filter had no significant effect on the numbers of patients showing improvement or progression. The use of 2 point change criteria significantly reduced the number of points progressing or improving compared to using a single point change criteria, regardless of whether the fields were blurred or not. **CONCLUSION**. Using criteria of 2 points with a slope >1 dB/year $P<0.05$ for detecting change is more likely to detect real progression in short field series. The use of a Gaussian filter is unhelpful.

Supported by: MRC grant G9330070, IGA.

Oxford Ophthalmological Congress, Oxford, 1998 - poster

The effect of number of visual fields and progression criteria on the detection of change in trabeculectomy patients

M.R.Wilkins, T.Lowe, A. Kotecha, R. Hitchings, P.T.Khaw, F.W.Fitzke

Wound Healing Research Unit and Department of Visual Sciences, Institute of Ophthalmology; Glaucoma Unit, Moorfields Eye Hospital, London, United Kingdom

We studied the effect of changing progression criteria on detecting worsening or improving visual fields following trabeculectomy. 48 patients listed for trabeculectomy had at least 2 Humphrey 24-2 fields prior to surgery followed by 4 fields at 3 month intervals. Field analysis was performed using pointwise linear regression over the second preoperative test and the 4 postoperative tests. Analysis was performed with and without a spatial filter applied to individual fields. Significant change in a point was defined as a slope of 1 dB/year $P<0.05$. Criteria for field series change was defined by the number and spatial arrangement of changing points. The results are shown in the table below.

Criteria for series change Number of changing points	Number of field series showing change					
	No Spatial Filter			Spatial Filter applied		
	Improving	Improving and progressing	Progressing	Improving	Improving and progressing	Progressing
1 point	12	12	13	12	7	16
Any 2 points	8	3	11	11	1	8
2 contiguous points	3	0	7	5	1	7

The criterion of 2 contiguous points reduces the number of field series changing, and reduces the number of fields showing both improvement and progression. With a single point criterion for series change spatial filtering reduces the number of series showing improvement and deterioration. Spatial filtering is of less obvious benefit where a 2-point change criterion is used.

Supported by: MRC grant G9330070, IGA.

The XIII International Perimetric Society Meeting, Lago di Garda, Italy, 1998 - poster

Pointwise linear regression criteria and the detection of change in automated visual field series

M.R.Wilkins, T.Lowe, A.Kotecha, R.Hitchings, P.T.Khaw, F.Fitzke

Wound Healing Research Unit, Department of Pathology, Institute of Ophthalmology; Glaucoma Unit, Moorfields Eye Hospital, London, United Kingdom

Sixty two patients undergoing trabeculectomy had at least 2 preoperative Humphrey 24-2 fields followed by 4 postoperative fields at 3 month intervals. Pointwise linear regression was used to analyse the field series; with the first field omitted from the analysis. Points were classified as deteriorating, or improving according to the magnitude and statistical significance of their slope. Either criterion was changed in small increments while the other was held constant. With the statistical significance held at $P<0.05$ and a critical slope of 0.01 dB/yr the numbers of patients with a single deteriorating, improving point or both were 34, 35, and 18 respectively. These figures fell to 21, 25, and 8 when the critical slope was increased to 5.0 dB/yr. With the critical slope held at 1.0 dB/yr and a statistical significance of $P<0.001$ the numbers of patients with a single deteriorating, improving point or both were 1, 0, and 0 respectively. These figures rose to 34, 35 and 18 when the statistical significance was set to $P<0.05$. Altering the critical slope within the range used is poor at separating field series showing deterioration from those showing improvement. Changing the statistical significance values within the range used offers better opportunities for separating deteriorating series from improving ones.

Supported by:MRC grant G9330070, IGA.

ARVO 1999, Fort Lauderdale - presentation

The effect of cluster progression criteria on the detection of field changes in trabeculectomy patients

M.R. Wilkins, A. Kotecha, D. Siriwardena, R. Hitchings, P.T. Khaw, F.W. Fitzke

Wound Healing Research Unit, Department of Pathology, Institute of Ophthalmology; Glaucoma Unit, Moorfields Eye Hospital, London, United Kingdom

PURPOSE We previously showed that by graphically displaying the number of patients changing as critical slope and P value were altered, optimal pointwise linear regression (PLR) criteria for detecting progression in a field series could be defined. This prevented detection of excessive change with less strict criteria. We therefore studied how criteria requiring a spatial relationship between points affects the detection of change in visual fields. **METHOD** We prospectively followed 56 patients undergoing trabeculectomy; analysing their Humphrey 24-2 visual fields using PLR. The last preoperative field plus 5 fields from the first 16 postoperative months were analysed. Points were labelled progressing or improving if they had a slope of 1 dB/yr and an optimal P value of <0.027. A chart plot of number of patients changing vs P value had shown this P value maximised detection of progression at the expense of improvement. Criteria for labelling a patient as changing were: 2 points changing anywhere in the field; 2 contiguous points in either vertical hemifield; and 2 points within a Glaucoma Hemifield Test (GHT) cluster. **RESULTS** No patients showed both progression and improvement. With a criterion of change at any 2 locations, 11 patients were progressing and 6 were improving. Requiring 2 contiguous points to change detected 5 patients progressing and 1 improving. Requiring 2 points to change in a GHT cluster detected 4 patients progressing and 0 improving. The 4 progressing patients were a subset of the 5 detected using the previous criterion. **CONCLUSION** Using spatial information to define change seems to selectively select progression in automated visual fields. In this series a progression criterion of 2 points in a GHT cluster changing with a slope of 1 dB/yr and with a P value of <0.027 exclusively selects progressing patients.

Supported by MRC (UK) grant G9330070.

ARVO 2001, Fort Lauderdale - presentation

Effect of stricter change criteria on the detection of change using Pointwise Linear Regression

Mark Wilkins, Roger Hitchings, Peng T Khaw, Fred Fitzke

Wound Healing Research Unit and Department of Visual Science - Institute of Ophthalmology. Glaucoma Unit Moorfields Eye Hospital.

PURPOSE To refine Pointwise Linear Regression (PLR) change criteria by varying P value, the number, and spatial arrangement of points. **METHOD** Having derived change criteria from an initial data set we selected optimal criteria and applied them to an independent set of 97 patients followed over 16 months. Criteria tested were: a single point changing anywhere $P<0.05$, $P<0.01$, $P<0.001$; 2 points changing anywhere $P<0.01$, 2 points in a Glaucoma Hemifield Test (GHT) cluster $P<0.01$, and 2 points in a Perimetric Nerve Fibre Bundle (PNFB) cluster $P<0.013$. Minimum slope was ± 1 dB/yr.

RESULTS

	Single point $P<0.05$	Single point $P<0.01$	Single point $P<0.001$	2 points anywhere $P<0.01$	2 points in a GHT cluster $P<0.01$	2 points in a PNFB cluster $P<0.013$
Patients progressing	65	25	7	7	1	7
Patients improving	45	15	2	3	1	0

CONCLUSION In developing change criteria for PLR, using a stricter P value does not enable detection of progression without some persisting detection of improvement. Using a criterion that specifies that points change in a PNFB cluster enables detection of progression without detecting improvement.

Supported by MRC (UK) grant G9330070.

AGIS (1994). "The Advanced Glaucoma Intervention Study (AGIS) 2. Visual field test scoring and reliability." Ophthalmology **101**(8): 1445-55.

AGIS (1994). "The Advanced Glaucoma Intervention Study (AGIS): 1. Study design and methods and baseline characteristics of study patients." Control Clin Trials **15**(4): 299-325.

AGIS (1998). "The Advanced Glaucoma Intervention Study (AGIS): 3. Baseline characteristics of black and white patients." Ophthalmology **105**(7): 1137-45.

AGIS (1998). "The Advanced Glaucoma Intervention Study (AGIS): 4. Comparison of treatment outcomes within race. Seven-year results." Ophthalmology **105**(7): 1146-64.

AGIS (2000). "The Advanced Glaucoma Intervention Study, 6: Effect of Cataract on Visual Field and Visual Acuity." Arch Ophthalmol **118**: 1639-1652.

Airaksinen, P. J., S. M. Drance, et al. (1984). "Diffuse and localized nerve fiber loss in glaucoma." American Journal of Ophthalmology **98**(5): 566-71.

Airaksinen, P. J., E. Mustonen, et al. (1981). "Optic disc haemorrhages precede retinal nerve fibre layer defects in ocular hypertension." Acta Ophthalmol Copenh **59**(5): 627-41.

Airaksinen, P. J., A. Tuulonen, et al. (1992). "Rate and pattern of neuroretinal rim area decrease in ocular hypertension and glaucoma." Arch Ophthalmol **110**(2): 206-10.

Altman, D. G. (1991). Practical Statistics for Medical research, Chapman & Hall.

Anctil, J. L. and D. R. Anderson (1984). "Early foveal involvement and generalized depression of the visual field in glaucoma." Arch Ophthalmol **102**(3): 363-70.

Anderson, D. R. and V. M. Patella (1999). Automated Static Perimetry, Mosby.

Anton, A., N. Yamagishi, et al. (1998). "Mapping structural to functional damage in glaucoma with standard automated perimetry and confocal scanning laser ophthalmoscopy." American Journal of Ophthalmology **125**(4): 436-46.

Anton, A., L. Zangwill, et al. (1997). "Nerve fiber layer measurements with scanning laser polarimetry in ocular hypertension." Archives of Ophthalmology **115**(3): 331-4.

Arden, G. B. and J. J. Jacobson (1978). "A simple grating test for contrast sensitivity: preliminary results indicate value in screening for glaucoma." Investigative Ophthalmology & Visual Science **17**: 23.

Arkell, S. M., D. A. Lightman, et al. (1987). "The prevalence of glaucoma among Eskimos of northwest Alaska." Archives of Ophthalmology **105**(4): 482-5.

Armaly, M. F. (1965). "On the distribution of applanation pressure.I.Statistical features and the effect of age, sex, and family history of glaucoma." Arch Ophthalmol **73**: 11.

Armaly, M. F., D. E. Krueger, et al. (1980). "Biostatistical analysis of the collaborative glaucoma study. I. Summary report of the risk factors for glaucomatous visual-field defects." Arch Ophthalmol **98**(12): 2163-71.

Asawaphureekorn, S., L. Zangwill, et al. (1996). "Ranked-segment distribution curve for interpretation of optic nerve topography." Journal of Glaucoma **5**(2): 79-90.

Åsman, P. and A. Heijl (1992). "Evaluation of methods for automated Hemifield analysis in perimetry." Arch Ophthalmol **110**(6): 820-6.

Åsman, P. and A. Heijl (1992). "Glaucoma Hemifield Test. Automated visual field evaluation." Arch Ophthalmol **110**(6): 812-9.

Åsman, P. and A. Heijl (1992). "Weighting according to location in computer-assisted glaucoma visual field analysis." Acta Ophthalmol Copenh **70**(5): 671-8.

Åsman, P. and A. Heijl (1994). "Diffuse visual field loss and glaucoma." Acta Ophthalmol Copenh **72**(3): 303-8.

Bailey, I. L. and J. E. Lovie (1976). "New design principles for visual acuity letter charts." Am J Optom Physiol Opt **53**(11): 740-5.

Balazsi, G. and E. B. Werner (1983). "Relationship between baring of circumlinear vessels of the optic disc and glaucomatous visual field loss." Canadian Journal of Ophthalmology **18**(7): 333-6.

Bankes, J. L., E. S. Perkins, et al. (1968). "Bedford glaucoma survey." British Medical Journal **1**(595): 791-6.

Bengtsson, B. and A. Heijl (1998). "Evaluation of a new perimetric threshold strategy, SITA, in patients with manifest and suspect glaucoma." Acta Ophthalmol Scand **76**(3): 268-72.

Bengtsson, B., A. Heijl, et al. (1998). "Evaluation of a new threshold visual field strategy, SITA, in normal subjects. Swedish Interactive Thresholding Algorithm." Acta Ophthalmologica Scandinavica **76**(2): 165-9.

Bengtsson, B., A. Lindgren, et al. (1997). "Perimetric probability maps to separate change caused by glaucoma from that caused by cataract." Acta Ophthalmologica Scandinavica **75**(2): 184-8.

Bengtsson, B., J. Olsson, et al. (1997). "A new generation of algorithms for computerized threshold perimetry, SITA." Acta Ophthalmologica Scandinavica **75**(4): 368-75.

Bhandari, A., D. P. Crabb, et al. (1997). "Effect of surgery on visual field progression in normal-tension glaucoma." Ophthalmology **104**(7): 1131-7.

Bickler-Bluth, M., G. L. Trick, et al. (1989). "Assessing the utility of reliability indices for automated visual fields. Testing ocular hypertensives [see comments]." Ophthalmology **96**(5): 616-9.

Birch, M. K., P. K. Wishart, et al. (1995). "Determining progressive visual field loss in serial Humphrey visual fields." Ophthalmology **102**(8): 1227-34; discussion 1234-5.

Boeglin, R. J., J. Caprioli, et al. (1992). "Long-term fluctuation of the visual field in glaucoma." Am J Ophthalmol **113**(4): 396-400.

Boles Carenini, A., G. Sibour, et al. (1994). "Differences in long term effect of timolol and betaxolol on the pulsatile ocular ocular blood flow." Survey of Ophthalmology **38**: 118-24.

Bonomi, L., G. Marchini, et al. (1997). "Prevalence of Glaucoma and Intraocular Pressure Distribution in a Defined Population. The Egna-Neumarkt Study." Ophthalmology **105**: 209-215.

Brigatti, L., D. Hoffman, et al. (1996). "Neural networks to identify glaucoma with structural and functional measurements." American Journal of Ophthalmology **121**(5): 511-21.

Britton, R. J., S. M. Drance, et al. (1987). "The area of the neuroretinal rim of the optic nerve in normal eyes." American Journal of Ophthalmology **103**(4): 497-504.

Brubaker, R. F. (1996). "Delayed functional loss in glaucoma. LII Edward Jackson Memorial Lecture." Am J Ophthalmol **121**(5): 473-83.

Budenz, D. L., W. J. Feuer, et al. (1993). "The effect of simulated cataract on the glaucomatous visual field." Ophthalmology **100**(4): 511-7.

Burke, J. and M. Schwartz (1996). "Preclinical evaluation of brimonidine." Survey of Ophthalmology **41**(Suppl 1): S9-18.

Cairns, J. E. (1969). "Trabeculectomy. Sub-title: "a surgical method of reducing intra-ocular pressure in chronic simple glaucoma without sub-conjunctival drainage of aqueous humour"." Trans Ophthalmol Soc U K **88**: 231-3.

Caprioli, J. (1992). "Discrimination between normal and glaucomatous eyes." Invest Ophthalmol Vis Sci **33**(1): 153-9.

Caprioli, J. and J. M. Miller (1987). "Optic disc rim area is related to disc size in normal subjects." Arch Ophthalmol **105**(12): 1683-5.

Caprioli, J., M. Sears, et al. (1987). "Patterns of early visual field loss in open-angle glaucoma." Am J Ophthalmol **103**(4): 512-7.

Casson, E. J., L. R. Shapiro, et al. (1990). "Short-term fluctuation as an estimate of variability in visual field data." Invest Ophthalmol Vis Sci **31**(11): 2459-63.

Chauhan, B. C., S. M. Drance, et al. (1990). "The use of visual field indices in detecting changes in the visual field in glaucoma." Investigative Ophthalmology & Visual Science **31**(3): 512-20.

Chauhan, B. C., R. P. LeBlanc, et al. (1997). "Repeatable diffuse visual field loss in open-angle glaucoma." Ophthalmology **104**(3): 532-8.

Chen, P. P. and D. L. Budenz (1998). "The effects of cataract extraction on the visual field of eyes with chronic open-angle glaucoma." American Journal of Ophthalmology **125**(3): 325-33.

Chylack, L. T., Jr., M. C. Leske, et al. (1989). "Lens opacities classification system II (LOCS II)." Archives of Ophthalmology **107**(7): 991-7.

Chylack, L. T., Jr., J. K. Wolfe, et al. (1993). "The Lens Opacities Classification System III. The Longitudinal Study of Cataract Study Group." Archives of Ophthalmology **111**(6): 831-6.

Clarke, M. P., S. A. Vernon, et al. (1990). "The development of cataract following trabeculectomy." Eye **4**(Pt 4): 577-83.

Coleman, A. L., H. A. Quigley, et al. (1991). "Displacement of the optic nerve head by acute changes in intraocular pressure in monkey eyes." Ophthalmology **98**(1): 35-40.

Collaborative Normal-Tension Glaucoma Study Group (1998). "Comparison of Glaucomatous Progression Between Untreated Patients With Normal-Tension Glaucoma and Patients With Therapeutically Reduced Pressures." AJO **126**(4): 487-497.

Collaborative Normal-Tension Glaucoma Study Group (1998). "The Effectiveness of Intraocular Pressure Reduction in the Treatment of Normal-Tension Glaucoma." AJO **126**(4): 498-505.

Colton, T. and F. Ederer (1980). "The Distribution of Intraocular Pressures in the General Population." Survey of Ophthalmology **25**: 123-129.

Corbett, J. J. and C. D. Phelps (1985). "The neurologic evaluation of patients with low tension glaucoma." Invest Ophthalmol Vis Sci **26**: 110-4.

Costa, V. P., M. Smith, et al. (1993). "Loss of visual acuity after trabeculectomy." Ophthalmology **100**(5): 599-612.

Costa, V. P., R. P. Wilson, et al. (1993). "Hypotony maculopathy following the use of topical mitomycin C in glaucoma filtration surgery." Ophthalmic Surg **24**(6): 389-94.

Crabb, D. P., D. F. Edgar, et al. (1995). "New approach to estimating variability in visual field data using an image processing technique." Br J Ophthalmol **79**(3): 213-7.

Crabb, D. P., F. W. Fitzke, et al. (1997). "Improving the prediction of visual field progression in glaucoma using spatial processing." Ophthalmology **104**(3): 517-24.

Crabb, D. P., F. W. Fitzke, et al. (1996). A Profile of the Spatial Dependence of Pointwise Sensitivity Across the Glaucomatous Visual Field. Perimetry Update 1996/1997, Wurzburg, Germany, Kugler Publications by Amsterdam/ New York.

Damms, T. and F. Dannheim (1993). "Sensitivity and specificity of optic disc parameters in chronic glaucoma." Invest Ophthalmol Vis Sci **34**(7): 2246-50.

David, R., D. G. Livingston, et al. (1977). "Ocular hypertension--a long-term follow-up of treated and untreated patients." British Journal of Ophthalmology **61**(11): 668-74.

DCCT (1998). "Early worsening of diabetic retinopathy in the Diabetes Control and Complications Trial." Arch Ophthalmol **116**(7): 874-86.

Desatnik, H., H. A. Quigley, et al. (1996). "Study of central retinal ganglion cell loss in experimental glaucoma in monkey eyes." Journal of Glaucoma **5**(1): 46-53.

Dielemans, I., J. R. Vingerling, et al. (1994). "The prevalence of primary open-angle glaucoma in a population-based study in The Netherlands. The Rotterdam Study." Ophthalmology **101**(11): 1851-5.

Drance, S. M. (1969). "The early field defects in glaucoma." Investigative Ophthalmology **8**(1): 84-91.

Drance, S. M. (1972). "Some factors in the production of low tension glaucoma." British Journal of Ophthalmology **56**(3): 229-42.

Drance, S. M. (1991). "Diffuse visual field loss in open-angle glaucoma." Ophthalmology **98**(10): 1533-8.

Drance, S. M. (1997). "Glaucoma: A look beyond intraocular pressure." Am J Ophthalmol **123**: 817-819.

Drance, S. M., V. Berry, et al. (1967). "Studies on the effects of age on the central and peripheral isopters of the visual field in normal subjects." American Journal of Ophthalmology **63**(6): 1667-72.

Duggan, C., A. Sommer, et al. (1985). "Automated differential threshold perimetry for detecting glaucomatous visual field loss." Am J Ophthalmol **100**(3): 420-3.

Egbert, P. R., A. S. Williams, et al. (1993). "A prospective trial of intraoperative fluorouracil during trabeculectomy in a black population." Am J Ophthalmol **116**(5): 612-6.

Epstein, D. L., J. H. Krug, Jr., et al. (1989). "A long-term clinical trial of timolol therapy versus no treatment in the management of glaucoma suspects [see comments]." Ophthalmology **96**(10): 1460-7.

ETDRS (1985). "Photocoagulation for diabetic macular edema. Early Treatment Diabetic Retinopathy Study report number 1. Early Treatment Diabetic Retinopathy Study research group." Arch Ophthalmol **103**(12): 1796-806.

ETDRS (1991). "Early Treatment Diabetic Retinopathy Study design and baseline patient characteristics. ETDRS report number 7." Ophthalmology **98**(5 Suppl): 741-56.

FFSG (1996). "Five-year follow-up of the Fluorouracil Filtering Surgery Study. The fluorouracil filtering surgery study group." Am J Ophthalmol **121**(4): 349-66.

Fitzke, F. W., D. P. Crabb, et al. (1995). "Image processing of computerised visual field data." Br J Ophthalmol **79**(3): 207-12.

Fitzke, F. W., R. A. Hitchings, et al. (1996). "Analysis of visual field progression in glaucoma." British Journal of Ophthalmology **80**(1): 40-8.

Fitzke, F. W. and A. I. McNaught (1994). "The diagnosis of visual field progression in glaucoma." Current Opinion in Ophthalmology **5**(11): 110-115.

Fitzke, F. W., D. Poinoosawmy, et al. (1989). Peripheral displacement thresholds in normals, ocular hypertensives and glaucoma. Seventh International Field Symposium, Amsterdam, Martinus Nijhoff/ Dr W.Junk.

Flammer, J., S. M. Drance, et al. (1985). "Quantification of glaucomatous visual field defects with automated perimetry." Invest Ophthalmol Vis Sci **26**(2): 176-81.

Flammer, J., S. M. Drance, et al. (1984). "Differential light threshold. Short- and long-term fluctuation in patients with glaucoma, normal controls, and patients with suspected glaucoma." Arch Ophthalmol **102**(5): 704-6.

Flanagan, J. G., I. D. Moss, et al. (1993). "Evaluation of FASTPAC: a new strategy for threshold estimation with the Humphrey Field Analyser." Graefes Arch Clin Exp Ophthalmol **231**(8): 465-9.

Flanagan, J. G., J. M. Wild, et al. (1993). "Evaluation of FASTPAC, a new strategy for threshold estimation with the Humphrey Field Analyzer, in a glaucomatous population." Ophthalmology **100**(6): 949-54.

Fleiss, J. L. (1981). Statistical Methods for Rates and Proportions. New York, Wiley.

Foster, P. J., J. Baasanhu, et al. (1996). "Glaucoma in Mongolia. A population-based survey in Hovsgol province, northern Mongolia." Archives of Ophthalmology **114**(10): 1235-41.

Funk, J. (1991). "Early detection of glaucoma by longitudinal monitoring of the optic disc structure." Graefes Arch Clin Exp Ophthalmol **229**(1): 57-61.

Gandolfi, S. A. (1995). "Improvement of visual field indices after surgical reduction of intraocular pressure." Ophthalmic Surg **26**(2): 121-6.

Garway-Heath, D. F., D. Poinoosawmy, et al. (2000). "Mapping the visual field to the optic disc in normal tension glaucoma eyes." Ophthalmology **107**(10): 1809-15.

Glaucoma Laser Trial Research Group (1991). "The Glaucoma Laser Trial (GLT): 3. Design and methods." Control Clin Trials **12**(4): 504-24.

Glaucoma Laser Trial Research Group (1995). "The Glaucoma Laser Trial (GLT) and glaucoma laser trial follow-up study: 7. Results." Am J Ophthalmol **120**(6): 718-31.

Glovinsky, Y., H. A. Quigley, et al. (1991). "Retinal ganglion cell loss is size dependent in experimental glaucoma." Investigative Ophthalmology & Visual Science **32**(3): 484-91.

Glovinsky, Y., H. A. Quigley, et al. (1993). "Foveal ganglion cell loss is size dependent in experimental glaucoma." Investigative Ophthalmology & Visual Science **34**(2): 395-400.

Greve, E. L. and C. L. Dake (1979). "Four year follow-up of a glaucoma operation. Prospective study of the double flap Scheie." Int Ophthalmol **1**(3): 139-45.

Gupta, N. and R. N. Weinreb (1997). "New definitions of glaucoma." Curr Op Ophthalmol **8**(II): 38-41.

Guthauser, U. and J. Flammer (1988). "Quantifying visual field damage caused by cataract." Am J Ophthalmol **106**(4): 480-4.

Haas, A., J. Flammer, et al. (1986). "Influence of age on the visual fields of normal subjects." Am J Ophthalmol **101**(2): 199-203.

Haefliger, I. O. and J. Flammer (1989). "Increase of the short-term fluctuation of the differential light threshold around a physiologic scotoma." Am J Ophthalmol **107**(4): 417-20.

Haefliger, I. O. and J. Flammer (1991). "Fluctuation of the differential light threshold at the border of absolute scotomas. Comparison between glaucomatous visual field defects and blind spots." Ophthalmology **98**(10): 1529-32.

Hart, W. M., Jr. and B. Becker (1982). "The onset and evolution of glaucomatous visual field defects." Ophthalmology **89**(3): 268-79.

Heijl, A. (1989). "Lack of diffuse loss of differential light sensitivity in early glaucoma." Acta Ophthalmol Copenh **67**(4): 353-60.

Heijl, A. and B. Bengtsson (1996). "The effect of perimetric experience in patients with glaucoma." Arch Ophthalmol **114**(1): 19-22.

Heijl, A., A. Lindgren, et al. (1989). "Test-retest variability in glaucomatous visual fields." Am J Ophthalmol **108**(2): 130-5.

Heijl, A., G. Lindgren, et al. (1990). Extended empirical statistical package for evaluation of single and multiple fields in glaucoma: Statpac 2. Perimetry Update 1990/91, Malmö, Sweden, Kugler Publications.

Heijl, A., G. Lindgren, et al. (1986). A package for the statistical analysis of visual fields. Documenta Ophthalmologica Proceedings. E. L. Greve and A. Heijl. Amsterdam, Martinus Nijhof/ Dr W. Junk. **49:** 153-168.

Heijl, A., G. Lindgren, et al. (1987). "Normal variability of static perimetric threshold values across the central visual field." Arch Ophthalmol **105**(11): 1544-9.

Heijl, A., G. Lindgren, et al. (1989). "The effect of perimetric experience in normal subjects." Archives of Ophthalmology **107**(1): 81-6.

Henson, D. B., P. H. Artes, et al. (1999). "Diffuse loss of sensitivity in early glaucoma." Investigative Ophthalmology & Visual Science **40**(13): 3147-51.

Henson, D. B., S. Chaudry, et al. (2000). "Response variability in the visual field: comparison of optic neuritis, glaucoma, ocular hypertension, and normal eyes." Investigative Ophthalmology & Visual Science **41**(2): 417-21.

Henson, D. B., S. E. Spenceley, et al. (1996). "Spatial classification of glaucomatous visual field loss." British Journal of Ophthalmology **80**(6): 526-31.

Heuer, D. K., D. R. Anderson, et al. (1987). "The influence of refraction accuracy on automated perimetric threshold measurements." Ophthalmology **94**(12): 1550-3.

Heuer, D. K., D. R. Anderson, et al. (1986). The influence of simulated media opacities on threshold measurements. Seventh International Visual Field Symposium, Amsterdam, Martinus Nijhoff/ Dr W. Junk.

Heuer, D. K., D. R. Anderson, et al. (1988). "The influence of simulated light scattering on automated perimetric threshold measurements." Arch Ophthalmol **106**(9): 1247-51.

Hitchings, R. A., C. S. Migdal, et al. (1994). "The primary treatment trial: changes in the visual field analysis by computer-assisted perimetry." Eye **8**(Pt 1): 117-20.

Holmin, C. and C. E. Krakau (1980). "Visual field decay in normal subjects and in cases of chronic glaucoma." Albrecht Von Graefes Arch Klin Exp Ophthalmol **213**(4): 291-8.

Holmin, C. and C. E. Krakau (1982). "Regression analysis of the central visual field in chronic glaucoma cases. A follow-up study using automatic perimetry." Acta Ophthalmol Copenh **60**(2): 267-74.

Hoskins, H. D., S. D. Magee, et al. (1988). "Confidence intervals for change in automated visual fields." Br J Ophthalmol **72**(8): 591-7.

Hoyt, W. F. and N. M. Newman (1972). "The earliest observable defect in glaucoma?" Lancet **1**(752): 692-3.

Jay, J. L. and D. Allan (1989). "The benefit of early trabeculectomy versus conventional management in primary open angle glaucoma relative to severity of disease." Eye **3**(Pt 5): 528-35.

Jay, J. L. and J. R. Murdoch (1993). "The rate of visual field loss in untreated primary open angle glaucoma." Br J Ophthalmol **77**(3): 176-8.

Johnson, C. A., A. J. Adams, et al. (1993). "Blue-on-yellow perimetry can predict the development of glaucomatous visual field loss." Arch Ophthalmol **111**(5): 645-50.

Johnson, C. A., A. J. Adams, et al. (1993). "Progression of early glaucomatous visual field loss as detected by blue-on-yellow and standard white-on-white automated perimetry." Arch Ophthalmol **111**(5): 651-6.

Johnson, C. A., A. J. Adams, et al. (1989). "Evidence for a neural basis of age-related visual field loss in normal observers." Invest Ophthalmol Vis Sci **30**(9): 2056-64.

Johnson, C. A., B. C. Chauhan, et al. (1992). "Properties of staircase procedures for estimating thresholds in automated perimetry." Invest Ophthalmol Vis Sci **33**(10): 2966-74.

Johnson, C. A., P. G. D. Spry, et al. (2000). "Evaluation of a variety of visual function test in ocular hypertension and glaucoma patients." Invest Ophthalmol Vis Sci **41**(4): S104.

Jonas, J. B., A. E. Grundler, et al. (1998). "Pressure-dependent neuroretinal rim loss in normal-pressure glaucoma." American Journal of Ophthalmology **125**(2): 137-44.

Jonas, J. B., A. M. Schmidt, et al. (1992). "Human optic nerve fiber count and optic disc size." Investigative Ophthalmology & Visual Science **33**(6): 2012-8.

Kahn, H. A., H. M. Leibowitz, et al. (1977). "The Framingham Eye Study. I. Outline and major prevalence findings." American Journal of Epidemiology **106**(1): 17-32.

Kaiser, H. J., J. Flammer, et al. (1994). "Longterm visual field follow-up of glaucoma patients treated with beta-blockers." Survey of Ophthalmology **38 Suppl**: S156-9; discussion S160.

Kamal, D. S., A. C. Viswanathan, et al. (1999). "Detection of optic disc change with the Heidelberg retina tomograph before confirmed visual field change in ocular hypertensives converting to early glaucoma." Br J Ophthalmol **83**(3): 290-4.

Kass, M. A., M. O. Gordon, et al. (1989). "Topical timolol administration reduces the incidence of glaucomatous damage in ocular hypertensive individuals. A randomized, double-masked, long-term clinical trial." Arch Ophthalmol **107**(11): 1590-8.

Katz, J. (1999). "Scoring systems for measuring progression of visual field loss in clinical trials of glaucoma treatment." Ophthalmology **106**(2): 391-5.

Katz, J., N. Congdon, et al. (1999). "Methodological Variations in Estimating Apparent Progressive Visual Field Loss in Clinical Trials of Glaucoma Treatment." Arch Ophthalmol **117**: 1137-1142.

Katz, J., D. Gilbert, et al. (1997). "Estimating progression of visual field loss in glaucoma." Ophthalmology **104**(6): 1017-25.

Katz, J., H. A. Quigley, et al. (1995). "Repeatability of the Glaucoma Hemifield Test in automated perimetry." Invest Ophthalmol Vis Sci **36**(8): 1658-24.

Katz, J., H. A. Quigley, et al. (1996). "Detection of incident field loss using the glaucoma hemifield test." Ophthalmology **103**(4): 657-63.

Katz, J. and A. Sommer (1988). "Reliability indexes of automated perimetric tests." Arch Ophthalmol **106**(9): 1252-4.

Katz, J. and A. Sommer (1990). "Screening for glaucomatous visual field loss. The effect of patient reliability." Ophthalmology **97**(8): 1032-7.

Katz, J., A. Sommer, et al. (1991). "Comparison of analytic algorithms for detecting glaucomatous visual field loss." Arch Ophthalmol **109**(12): 1684-9.

Katz, J., A. Sommer, et al. (1991). "Reliability of visual field results over repeated testing." Ophthalmology **98**(1): 70-5.

Katz, J., J. M. Tielsch, et al. (1995). "Automated perimetry detects visual field loss before manual Goldmann perimetry." Ophthalmology **102**(1): 21-6.

Keltner, J. L., C. A. Johnson, et al. (1993). "Baseline visual field profile of optic neuritis. The experience of the optic neuritis treatment trial. Optic Neuritis Study Group [see comments]." Arch Ophthalmol **111**(2): 231-4.

Keltner, J. L., C. A. Johnson, et al. (1994). "Visual field profile of optic neuritis. One-year follow-up in the Optic Neuritis Treatment Trial." Arch Ophthalmol **112**(7): 946-53.

Kendell, K. R., H. A. Quigley, et al. (1995). "Primary open-angle glaucoma is not associated with photoreceptor loss." Invest Ophthalmol Vis Sci **36**(1): 200-5.

Kidd, M. N. and O. C. M (1985). "Progression of field loss after trabeculectomy: a five-year follow-up." Br J Ophthalmol **69**(11): 827-31.

King, A. J., D. Farnworth, et al. (1997). "Inter-observer and intra-observer agreement in the interpretation of visual fields in glaucoma." Eye **11**(Pt 5): 687-91.

Kitazawa, Y., T. Horie, et al. (1977). "Untreated ocular hypertension. A long-term prospective study." Arch Ophthalmol **95**(7): 1180-4.

Klein, B. E., R. Klein, et al. (1992). "Prevalence of glaucoma. The Beaver Dam Eye Study." Ophthalmology **99**(10): 1499-504.

Kolker, A. E. (1977). "Visual prognosis in advanced glaucoma: a comparison of medical and surgical therapy for retention of vision in 101 eyes with advanced glaucoma." Trans Am Ophthalmol Soc **75**: 539-55.

Lachenmayr, B. J., S. M. Drance, et al. (1991). "Diffuse and localized glaucomatous field loss in light-sense, flicker and resolution perimetry." Graefes Archive for Clinical & Experimental Ophthalmology **229**(3): 267-73.

Lachenmayr, B. J., S. M. Drance, et al. (1991). "Light-sense, flicker and resolution perimetry in glaucoma: a comparative study." Graefes Archive for Clinical & Experimental Ophthalmology **229**(3): 246-51.

Lachenmayr, B. J., U. Kiermeir, et al. (1995). "Points of a normal visual field are not statistically independent." Ger J Ophthalmol **4**(3): 175-81.

Lam, B. L., W. L. Alward, et al. (1991). "Effect of cataract on automated perimetry." Ophthalmology **98**(7): 1066-70.

Lanigan, L., J. Sturmer, et al. (1994). "Single intraoperative applications of 5-fluorouracil during filtration surgery: early results." Br J Ophthalmol **78**(1): 33-7.

Leske, M. C., A. M. Connell, et al. (1995). "Risk factors for open-angle glaucoma. The Barbados Eye Study." Arch Ophthalmol **113**(7): 918-24.

Leske, M. C., A. M. Connell, et al. (2001). "Incidence of open-angle glaucoma: the Barbados Eye Studies. The Barbados Eye Studies Group." Archives of Ophthalmology **119**(1): 89-95.

Leske, M. C., A. Heijl, et al. (1999). "Early Manifest Glaucoma Trial: design and baseline data." Ophthalmology **106**(11): 2144-53.

Lewis, R. A., C. A. Johnson, et al. (1986). "Variability of quantitative automated perimetry in normal observers." Ophthalmology **93**(7): 878-81.

Lichter, P. R. (1976). "Variability of expert observers in evaluating the optic disc." Trans Am Ophthalmol Soc **84**: 532.

Lindenmuth, K. A., G. L. Skuta, et al. (1989). "Effects of pupillary constriction on automated perimetry in normal eyes." Ophthalmology **96**(9): 1298-301.

Liu, S., S. V. Araujo, et al. (1996). "Lack of effect of calcium channel blockers on open-angle glaucoma." Journal of Glaucoma **5**(3): 187-90.

Mao, L. K., W. C. Stewart, et al. (1991). "Correlation between intraocular pressure control and progressive glaucomatous damage in primary open-angle glaucoma." Am J Ophthalmol **111**(1): 51-5.

McNaught, A. I., D. P. Crabb, et al. (1996). "Visual field progression: comparison of Humphrey Statpac2 and pointwise linear regression analysis." Graefes Archive for Clinical & Experimental Ophthalmology **234**(7): 411-8.

Migdal, C., W. Gregory, et al. (1994). "Long-term functional outcome after early surgery compared with laser and medicine in open-angle glaucoma." Ophthalmology **101**(10): 1651-6.

Mikelberg, F. S. and S. M. Drance (1984). "The mode of progression of visual field defects in glaucoma." American Journal of Ophthalmology **98**(4): 443-5.

Miller, K. M. and H. A. Quigley (1988). "The clinical appearance of the lamina cribrosa as a function of the extent of glaucomatous optic nerve damage." Ophthalmology **95**(1): 135-8.

Motolko, M. and S. M. Drance (1981). "Features of the optic disc in preglaucomatous eyes." Archives of Ophthalmology **99**(11): 1992-4.

MPS (1982). "Argon laser photocoagulation for senile macular degeneration. Results of a randomized clinical trial." Arch Ophthalmol **100**(6): 912-8.

Netland, P. A., N. Chaturvedi, et al. (1993). "Calcium channel blockers in the management of low-tension and open-angle glaucoma." Am J Ophthalmol **115**(5): 608-13.

Noureddin, B. N., D. Poinoosawmy, et al. (1991). "Regression analysis of visual field progression in low tension glaucoma." Br J Ophthalmol **75**(8): 493-5.

Nouri-Mahdavi, K., L. Brigatti, et al. (1995). "Outcomes of trabeculectomy for primary open-angle glaucoma." Ophthalmology **102**(12): 1760-9.

Nouri-Mahdavi, K., L. Brigatti, et al. (1997). "Comparison of methods to detect visual field progression in glaucoma." Ophthalmology **104**(8): 1228-36.

O' Brien, C. and B. Schwartz (1990). "The visual field in chronic open angle glaucoma: the rate of change in different regions of the field." Eye **4**(Pt 4): 557-62.

O' Brien, C., B. Schwartz, et al. (1991). "Intraocular pressure and the rate of visual field loss in chronic open-angle glaucoma." Am J Ophthalmol **111**(4): 491-500.

O' Connor, D. J., T. Zeyen, et al. (1993). "Comparisons of methods to detect glaucomatous optic nerve damage." Ophthalmology **100**(10): 1498-503.

Odberg, T. (1993). "Visual field prognosis in early glaucoma. A long-term clinical follow-up." Acta Ophthalmol Copenh **71**(6): 721-6.

ONTT (1991). "The clinical profile of optic neuritis. Experience of the Optic Neuritis Treatment Trial. Optic Neuritis Study Group." Arch Ophthalmol **109**(12): 1673-8.

Osborne, N. N., C. Cazevieille, et al. (1997). "In vivo and in vitro experiments show that betaxolol is a retinal neuroprotective agent." Brain Research **751**(1): 113-23.

Palmberg, P. (1996). Preferred practice pattern for primary open-angle glaucoma. San Francisco, American Academy of Ophthalmology.

Panda, S. and J. B. Jonas (1992). "Decreased photoreceptor count in human eyes with secondary angle-closure glaucoma." Investigative Ophthalmology & Visual Science **33**: 2532-2536.

Pearson, P. A., R. S. Baker, et al. (1991). "Evaluation of optic nerve sheath fenestration in pseudotumor cerebri using automated perimetry." Ophthalmology **98**(1): 99-105.

Phelps, C. D. and J. J. Corbett (1985). "Migraine and low-tension glaucoma. A case control study." Invest Ophthal Vis Sci **26**: 1105-8.

Phillips, C. I., R. L. Vaid, et al. (1978). "Field defect due to posterior cortical paraxial lens opacity. A case report." Transactions of the Ophthalmological Societies of the United Kingdom **98**(4): 486-9.

Piltz, J. R., S. Bose, et al. (1998). "The Effect of Nimodipine, a Centrally Active Calcium Channel Antagonist, on Visual Function and Macular Blood Flow in Patients with Normal-Tension Glaucoma ad Control Subjects." J.Glaucoma **7**: 336-342.

Quigley, H. A. (1996). "Number of people with glaucoma worldwide." British Journal of Ophthalmology **80**(5): 389-93.

Quigley, H. A. and E. M. Addicks (1981). "Regional differences in the structure of the lamina cribrosa and their relation to glaucomatous optic nerve damage." Archives of Ophthalmology **99**(1): 137-43.

Quigley, H. A., E. M. Addicks, et al. (1982). "Optic nerve damage in human glaucoma. III. Quantitative correlation of nerve fiber loss and visual field defect in glaucoma, ischemic neuropathy, papilledema, and toxic neuropathy." Arch Ophthalmol **100**(1): 135-46.

Quigley, H. A., A. L. Coleman, et al. (1991). "Larger optic nerve heads have more nerve fibers in normal monkey eyes." Archives of Ophthalmology **109**(10): 1441-3.

Quigley, H. A., G. R. Dunkelberger, et al. (1988). "Chronic human glaucoma causing selectively greater loss of large optic nerve fibers." Ophthalmology **95**(3): 357-63.

Quigley, H. A., G. R. Dunkelberger, et al. (1989). "Retinal ganglion cell atrophy correlated with automated perimetry in human eyes with glaucoma." Am J Ophthalmol **107**(5): 453-64.

Quigley, H. A., R. M. Hohman, et al. (1983). "Morphologic changes in the lamina cribrosa correlated with neural loss in open-angle glaucoma." Am J Ophthalmol **95**(5): 673-91.

Quigley, H. A., J. Katz, et al. (1992). "An evaluation of optic disc and nerve fiber layer examinations in monitoring progression of early glaucoma damage." Ophthalmology **99**(1): 19-28.

Quigley, H. A., R. M. Sanchez, et al. (1987). "Chronic glaucoma selectively damages large optic nerve fibers." Investigative Ophthalmology & Visual Science **28**(6): 913-20.

Rebolleda, G., F. J. Munoz, et al. (1992). "Effects of pupillary dilation on automated perimetry in glaucoma patients receiving pilocarpine." Ophthalmology **99**(3): 418-23.

Robin, A. L., R. Ramakrishnan, et al. (1997). "A long-term dose-response study of mitomycin in glaucoma filtration surgery." Archives of Ophthalmology **115**(8): 969-74.

Rock, W. J., S. M. Drance, et al. (1971). "A modification of the Armaly visual field screening technique for glaucoma." Canadian Journal of Ophthalmology **6**(4): 283-92.

Rock, W. J., S. M. Drance, et al. (1973). "Visual field screening in glaucoma. An evaluation of the Armaly technique for screening glaucomatous visual fields." Archives of Ophthalmology **89**(4): 287-90.

Rollins, D. F. and S. M. Drance (1981). "Five year follow-up of trabeculectomy in the management of

chronic open angle glaucoma." Trans New Orleans Acad Ophtahlmol: 295-300.

Rossetti, L., I. Marchetti, et al. (1993). "Randomized clinical trials on medical treatment of glaucoma. Are they appropriate to guide clinical practice?" Arch Ophthalmol **111**(1): 96-103.

Sawada, A., Y. Kitazawa, et al. (1996). "Prevention of visual field defect progression with brovincamine in eyes with normal-tension glaucoma." Ophthalmology **103**(2): 283-8.

Schultz, J. S., E. B. Werner, et al. (1987). "Intraocular pressure and visual field defects after argon laser trabeculoplasty in chronic open-angle glaucoma." Ophthalmology **94**(5): 553-7.

Schulzer, M. (1994). "Errors in the diagnosis of visual field progression in normal-tension glaucoma." Ophthalmology **101**(9): 1589-94; discussion 1595.

Schulzer, M., D. R. Anderson, et al. (1991). "Sensitivity and specificity of a diagnostic test determined by repeated observations in the absence of an external standard." Journal of Clinical Epidemiology **44**(11): 1167-79.

Schulzer, M., S. M. Drance, et al. (1991). "A comparison of treated and untreated glaucoma suspects." Ophthalmology **98**(3): 301-7.

Smith, R. J. (1986). "The Lang lecture 1986. The enigma of primary open-angle glaucoma." Trans Ophthalmol Soc U K **105**(Pt 6): 618-33.

Smith, S. D., J. Katz, et al. (1996). "Analysis of progressive change in automated visual fields in glaucoma." Invest Ophthalmol Vis Sci **37**(7): 1419-28.

Smith, S. D., J. Katz, et al. (1997). "Effect of cataract extraction on the results of automated perimetry in glaucoma." Archives of Ophthalmology **115**(12): 1515-9.

Sommer, A. (1989). "Intraocular pressure and glaucoma." American Journal of Ophthalmology **107**(2): 186-8.

Sommer, A., C. Enger, et al. (1987). "Screening for glaucomatous visual field loss with automated threshold perimetry." Am J Ophthalmol **103**(5): 681-4.

Sommer, A., I. Pollack, et al. (1979). "Optic disc parameters and onset of glaucomatous field loss. I. Methods and progressive changes in disc morphology." Arch Ophthalmol **97**(8): 1444-8.

Sommer, A., J. M. Tielsch, et al. (1991). "Relationship between intraocular pressure and primary open angle glaucoma among white and black Americans. The Baltimore Eye Survey." Archives of Ophthalmology **109**(8): 1090-5.

Sponsel, W. E., N. L. Dallas, et al. (1983). "Visual field survival: the response to timolol therapy in open-angle glaucoma." Br J Ophthalmol **67**(4): 220-7.

Stamper, R. L., M. G. McMenemy, et al. (1992). "Hypotonous maculopathy after trabeculectomy with subconjunctival 5-fluorouracil " Am J Ophthalmol **114**(5): 544-53.

Stewart, W. C., G. M. Rogers, et al. (1995). "Effect of cataract extraction on automated fields in chronic open-angle glaucoma." Arch Ophthalmol **113**(7): 875-9.

Thomas, R., T. Kuriakose, et al. (1996). "Glaucoma like defect on automated perimetry caused by cataract." Indian Journal of Ophthalmology **44**(3): 167-8.

Tielsch, J. M., J. Katz, et al. (1991). "A population-based evaluation of glaucoma screening: the Baltimore Eye Survey." American Journal of Epidemiology **134**(10): 1102-10.

Tuulonen, A. and P. J. Airaksinen (1991). "Initial glaucomatous optic disk and retinal nerve fiber layer abnormalities and their progression." American Journal of Ophthalmology **111**(4): 485-90.

Tuulonen, A., J. Lehtola, et al. (1993). "Nerve fiber layer defects with normal visual fields. Do normal optic disc and normal visual field indicate absence of glaucomatous abnormality?" Ophthalmology **100**(5): 587-97.

Van Coevorden, R. E., R. P. Mills, et al. (1999). "Continuous Visual Field Test Supervision May Not ALways Be Necessary." Ophthalmology **106**: 178-181.

Varma, R., S. C. Hilton, et al. (1995). "Neural rim area declines with increased intraocular pressure in urban Americans [see comments]." Arch Ophthalmol **113**(8): 1001-5.

Varma, R., G. L. Spaeth, et al. (1989). "Agreement between clinicians and an image analyzer in estimating cup-to-disc ratios." Archives of Ophthalmology **107**(4): 526-9.

Varma, R., W. C. Steinmann, et al. (1992). "Expert agreement in evaluating the optic disc for glaucoma." Ophthalmology **99**(2): 215-21.

Varma, R., J. M. Tielsch, et al. (1994). "Race-, age-, gender-, and refractive error-related differences in the normal optic disc." Arch Ophthalmol **112**(8): 1068-76.

Viswanathan, A. C., F. W. Fitzke, et al. (1996). Spatial Filtering of Glaucomatous Visual Fields Using Progressor For Windows. Perimetry Update 1996/1997, Wurzburg, Germany, Kugler Publications by Amsterdam/ New York.

Viswanathan, A. C., F. W. Fitzke, et al. (1997). "Early detection of visual field progression in glaucoma: a comparison of PROGRESSOR and STATPAC 2." Br J Ophthalmol **81**: 1037-1042.

Wall, M., C. A. Johnson, et al. (1998). "Long- and short-term variability of automated perimetry results in patients with optic neuritis and healthy subjects." Archives of Ophthalmology **116**(1): 53-61.

Waring, G. O. d., S. D. Moffitt, et al. (1983). "Rationale for and design of the National Eye Institute Prospective Evaluation of Radial Keratotomy (PERK) Study." Ophthalmology **90**(1): 40-58.

Weber, J., F. Dannheim, et al. (1990). "The topographical relationship between optic disc and visual field in glaucoma." Acta Ophthalmol Copenh **68**(5): 568-74.

Weber, J. and H. Ulrich (1991). "A perimetric nerve fiber bundle map." Int Ophthalmol **15**(3): 193-200.

Webster, A. R., A. J. Luff, et al. (1993). "The effect of pilocarpine on the glaucomatous visual field." Br J Ophthalmol **77**(11): 721-5.

Weinreb, R. N. and J. P. Perlman (1986). "The effect of refractive correction on automated perimetric thresholds." Am J Ophthalmol **101**(6): 706-9.

Werner, E. B., A. Adelson, et al. (1988). "Effect of patient experience on the results of automated perimetry in clinically stable glaucoma patients." Ophthalmology **95**(6): 764-7.

Werner, E. B., K. I. Bishop, et al. (1988). "A comparison of experienced clinical observers and statistical tests in detection of progressive visual field loss in glaucoma using automated perimetry." Arch Ophthalmol **106**(5): 619-23.

Werner, E. B. and S. M. Drance (1977). "Early visual field disturbances in glaucoma." Arch Ophthalmol **95**(7): 1173-5.

Werner, E. B., S. M. Drance, et al. (1977). "Trabeculectomy and the progression of glaucomatous visual field loss." Arch Ophthalmol **95**(8): 1374-7.

Werner, E. B., B. Petrig, et al. (1989). "Variability of automated visual fields in clinically stable glaucoma patients." Invest Ophthalmol Vis Sci **30**(6): 1083-9.

Werner, E. B., N. Saheb, et al. (1982). "Lack of generalized constriction of affected visual field in glaucoma patients with visual field defects in one eye." Canadian Journal of Ophthalmology **17**(2): 53-5.

Werner, E. B., N. Saheb, et al. (1982). "Variability of static visual threshold responses in patients with elevated IOPs." Arch Ophthalmol **100**(10): 1627-31.

Wild, J. M., M. Dengler Harles, et al. (1989). "The influence of the learning effect on automated perimetry in patients with suspected glaucoma." Acta Ophthalmol Copenh **67**(5): 537-45.

Wild, J. M., M. K. Hussey, et al. (1993). "Pointwise topographical and longitudinal modeling of the visual field in glaucoma." Invest Ophthalmol Vis Sci **34**(6): 1907-16.

Wild, J. M., N. Hutchings, et al. (1997). "Pointwise Univariate Linear Regression of Periemtric Sensitivity Against Follow-up Time in Glaucoma." Ophthalmology **104**: 808-815.

Wild, J. M., A. E. Searle, et al. (1991). "Long-term follow-up of baseline learning and fatigue effects in the automated perimetry of glaucoma and ocular hypertensive patients." Acta Ophthalmol Copenh **69**(2): 210-6.

Wollstein, G., D. F. Garway-Heath, et al. (1998). "Identification of early glaucoma cases with the scanning laser ophthalmoscope." Ophthalmology **105**(8): 1557-63.

Wood, J. M., J. M. Wild, et al. (1988). "Factors affecting the normal perimetric profile derived by automated static threshold LED perimetry. I. Pupil size." Ophthalmic & Physiological Optics **8**(1): 26-31.

Wood, J. M., J. M. Wild, et al. (1987). "Serial examination of the normal visual field using Octopus automated projection perimetry. Evidence for a learning effect." Acta Ophthalmol Copenh **65**(3): 326-33.

Wood, J. M., J. M. Wild, et al. (1989). "Alterations in the shape of the automated perimetric profile arising from cataract." Graefes Archive for Clinical & Experimental Ophthalmology **227**(2): 157-61.

Wu, D.-C., B. Schwartz, et al. (1986). Trend analyses of automated visual fields. Seventh International Visual Field Symposium, Amsterdam, Martinus Nijhoff/ Dr W. Junk.

Wygnanski, T., H. Desatnik, et al. (1995). "Comparison of ganglion cell loss and cone loss in experimental glaucoma." American Journal of Ophthalmology **120**(2): 184-9.

Yamagishi, N., A. Anton, et al. (1997). "Mapping structural damage of the optic disk to visual field defect in glaucoma." American Journal of Ophthalmology **123**(5): 667-76.

Yoles, E. and M. Schwartz (1998). "Potential Neuroprotective Therapy for Glaucomatous Optic Neuropathy." Survey of Ophthalmology **42**: 367-72.

Yoshiyama, K. K. and C. A. Johnson (1997). "Which method of flicker perimetry is most effective for detection of glaucomatous visual field loss?" Invest Ophthalmol Vis Sci **38**(11): 2270-7.

Zeyen, T. G. and J. Caprioli (1993). "Progression of disc and field damage in early glaucoma." Arch Ophthalmol **111**(1): 62-5.