Studies on real world visual field data in glaucoma

Luke John Saunders

A thesis submitted for the degree of Doctor of Philosophy

CITY UNIVERSITY LONDON

School of Health Sciences

February 2015

City, University of London
Northampton Square
London
EC1V 0HB
United Kingdom

T +44 (0)20 7040 5060

THE FOLLOWING PART OF THIS THESIS HAS BEEN PUBLISHED:

Work in Chapter 4 has been published as the following article:
Saunders, L.J., Russell, R.A. and Crabb, D.P., (2015). Measurement Precision in a Series of Visual Fields Acquired by the Standard and Fast Versions of the Swedish Interactive Thresholding Algorithm: Analysis of Large-Scale Data From Clinics, *JAMA Ophthalmology,* **133**(1): 74-80.

doi:10.1001/jamaophthalmol.2014.4237

Table of Contents

List of Tables

List of Figures

Acknowledgements

First and most importantly of all, I would like to thank my supervisors Professor David Crabb and Dr Richard Russell for all their support over the last three years, not least for their involvement in all of the projects contained within this thesis. David's ideas, ambition and encouragement have been invaluable in my progression as a researcher, whilst Richard has been as much an invaluable friend as well as a supervisor and has always been around when I have needed guidance and advice. I am so grateful and indebted to them both.

I am also much obliged for the support that I have received from Moorfields Eye Hospital and for all the clinical data used in the study. In particular, I would like to thank Ted Garway-Heath an am hugely grateful for the opportunity to present my research at Moorfields Glaucoma Seminars. Andrew McNaught (Department of Ophthalmology, Gloucestershire Hospitals NHS Foundation Trust, Cheltenham and Cranfield University, Bedford), James Kirwan (Department of Ophthalmology, Queen Alexandra Hospital, Portsmouth) and Nitin Anand (Calderdale and Huddersfield NHS Foundation Trust) further deserve acknowledgement for providing access to visual field data from their respective hospitals.

This thesis would not have been possible without the funding I have received from my City University of London studentship, which has also been co-funded by an unrestricted grant from Allergan Inc. I am also highly grateful to the University for giving me so many opportunities to present my research and allowing me to receive feedback that has greatly enhanced my presentation skills.

It is finally important to acknowledge all the people who have supported me over the past three years, especially all my colleagues in the Crabb Lab. It has been a massive pleasure working alongside them and they have provided me with friendship, encouragement, comments, suggestions and useful advice. I am much obliged to Angharad Hobby, in particular, who helped proof the background sections of my thesis to ensure accuracy from a clinical perspective. I would further like to thank my wife, Diana, my mother and my sister for aiding me in the

proof reading of my work and for generally being an amazing support for me through the ups and downs of my three years at City University. I would like to also mention how grateful I am to my parents and grandparents for the financial support and encouragement for initially re-training in Statistics. Without this, there is no chance I would have even made it to City University to write this thesis, so I am incredibly fortunate to even have this opportunity. Finally, a massive thanks all my family, my amazing friends at the Journey church and beyond and, most of all, to God. I am so blessed by you all.

Declaration

The work contained in this thesis was completed by the candidate, Luke John Saunders. It has not been submitted for any other degrees, either now or in the past.

Where work contained within it has been previously published, this has been stated in the text. All sources of information have been acknowledged and references have been given.

The University Librarian of City University London is permitted to allow the thesis to be copied in whole or in part without further reference to the author. This permission covers only single copies made for study purposes, subject to normal conditions of acknowledgement.

Abstract

Glaucoma is a leading cause of blindness. As a progressive condition, it is important to monitor how the visual field (VF) changes over time with perimetry in preventing vision from deteriorating to a stage where quality of life is affected. However, there is little evidence of how clinical measurements correlate with meaningful quality of life landmarks for the patient or, by extension, the proportion of patients in danger of progressing to these landmarks. Further, measurement variability associated with visual fields make it difficult to monitor true change over time. The purpose of this thesis was to use large-scale clinical data (almost 500,000 VFs) to address some of these issues.

The first study attempted to relate clinical measurements of glaucoma severity to UK legal fitness to drive status. Legal fitness to drive (LFTD) was estimated using the integrated visual field as a surrogate of the Esterman test, which is the approved method by the UK DVLA of defining LFTD, while the mean deviation (MD) was used to represent defect severity. An MD of -14dB or worse in the better eye was found to be associated with a 92% (95% Confidence Interval [CI]: 87-95%) probability of being legally unfit to drive.

The second study used a statistical model to estimate the number of patients progressing at rates that could lead to this landmark of significant visual impairment or blindness in their predicted remaining lifetime. A significant minority of patients were progressing at rates that could lead to statutory blindness, as defined by the US Social Security Administration, in their predicted remaining lifetime (5.2% [CI: 4.5-6.0%]) with a further 10% in danger of becoming legally unfit to drive (10.4% [CI: 9.4-11.4%]). More than 90% (CI: 85.7-94.3%) of patients predicted to progress to statutory blindness had an MD worse than -6dB in at least one eye at presentation, suggesting an association between baseline VF damage and risk of future impairment.

The next section investigated whether choice of testing algorithm, SITA Standard or SITA Fast, affected the time taken to detect progression in VF follow-up. The precision of the tests was measured using linear modelling techniques and the impact of these differences was analysed using simulations. Though SITA Fast was found to be slightly less precise, no evidence was found to suggest that this resulted in progression being detected later.

The final study evaluated a validated and published risk calculator, which utilised baseline risk factors to profile risk of fast progression. A simpler model using baseline VF data was developed to have similar statistical properties for comparison (including equivalent R^2 statistics). The results suggested that risk calculators with low R^2 statistics had little utility for predicting future progression rate in clinical practice.

Together these results contribute a variety of novel findings and demonstrate the benefit of using large quantities of data collected from the everyday clinical milieu to extend clinical knowledge.

List of Abbreviations and Terms

ACG	Angle closure glaucoma
ADREV	Assessment of Disability Related to Vision
AGIS	Advanced Glaucoma Intervention Study
AIGS	Advanced Imaging in Glaucoma Study
ANSWERS	Analysis with Non-Stationary Weibull Error Regression with spatial enhancement
AUC	Area Under the curve
Beta-PPA	Beta-zone Parapapillary Atrophy
BEMD	Better Eye Mean Deviation
CCT	Central Corneal Thickness
CGS	Canadian Glaucoma Study
CI	Confidence Interval
CIGTS	Collaborative Initial Glaucoma Treatment Study
CNTGS	Collaborative Normal Tension Glaucoma Study
CPSD	Corrected Pattern Standard Deviation
DH	Disc Haemorrhage
DVLA	Driving and Vehicle Licensing Agency
EMGT	Early Manifest Glaucoma Trial
ERF	Error Related Factor
FDP	Frequency Doubling Perimetry
FDT	Frequency Doubling Technology
FL	Fixation losses
FN	False negative
FP	False positive
GCP	Glaucoma Change Probability
GSS	Glaucoma Staging System
H-P-A	Hodapp-Parrish-Anderson

HFA	Humphrey Field Analyzer
IOP	Intraocular Pressure
IVF	Integrated Visual Field
IQR	Interquartile Range
LFTD	Legal Fitness to Drive
LFTDP	Legally Fit to Drive Patients
LUTDP	Legally Unfit to Drive Patients
MD	Mean Deviation
NICE	National Institute for Health and Clinical Excellence
NPV	Negative Predictive Value
NTG	Normal Tension Glaucoma
NY-GAPS	New York Glaucoma Progression Study
OHT	Ocular Hypertension
OHTS	Ocular Hypertension Treatment Study
OLSR	Ordinary Least Squares Regression
ONH	Optic Nerve Head
ONS	Office of National Statistics
OPP	Ocular Perfusion Pressure
PoF	Probability of failure (the positive predictive value)
PD	Pattern deviation
PLR	Pointwise Linear Regression
POAG	Primary open angle glaucoma
PSD	Pattern Standard deviation
POAG	Primary Open Angle Glaucoma
PROM	Patient Reported Outcome Measure
QoL	Quality of Life
ROC	Receiver Operating Characteristic
SAP	Standard Automated Perimetry

SITA	Swedish Interactive Thresholding Algorithm
SSI	Severely Sight Impaired
SWAP	Short-wavelength automated perimetry
TD	Total deviation
UKGTS	United Kingdom Glaucoma Treatment Study
USP-GVFSS	University of Sao Paulo Glaucoma Visual Field Staging System
VF	Visual Field
VFI	Visual Field Index
WEMD	Worse Eye Mean Deviation
ZATA	Zippy Adaptive Threshold Algorithm

Chapter One: Background and Aims

This introductory chapter sets out to briefly go over the important background information underpinning my topic and defining the research questions that I have set out to answer in this thesis. It begins by briefly describing what glaucoma is, the risk factors associated with its incidence and progression and the way that a patient visual field (VF; this refers to the full extent of what an eye can see) is measured. The importance of monitoring loss effectively over time, the means of doing so and the problems associated with this VF loss will also be looked at, thus, setting the groundwork necessary to introduce how the work in this thesis contributes to current clinical understanding.

1.1 Glaucoma

Glaucoma is a group of optic neuropathies in which the optic nerve head and retinal ganglion cells are damaged potentially causing blindness. The eye disease is the second leading cause of blindness globally (**Figure 1.1**) and the leading cause of irreversible blindness worldwide affecting an estimated 60.5 million people worldwide with 8.4 million blind from the disease (Quigley & Broman 2006, World Health Organisation 2007, National Eye Institute 2010). In the UK, glaucoma is the main attributed cause of 10% of the cases of blindness (National Institute for Health and Clinical Excellence 2009). Treatment and monitoring of the condition is behind over one million hospital visits each year (National Institute for Health and Clinical Excellence 2009) and, due to the fact that the condition becomes more prevalent in elderly populations (Khawaja et al. 2013), it represents an even larger challenge to resources and healthcare in the future as global life expectancies increase (Quigley & Broman 2006, National Eye Institute 2010).

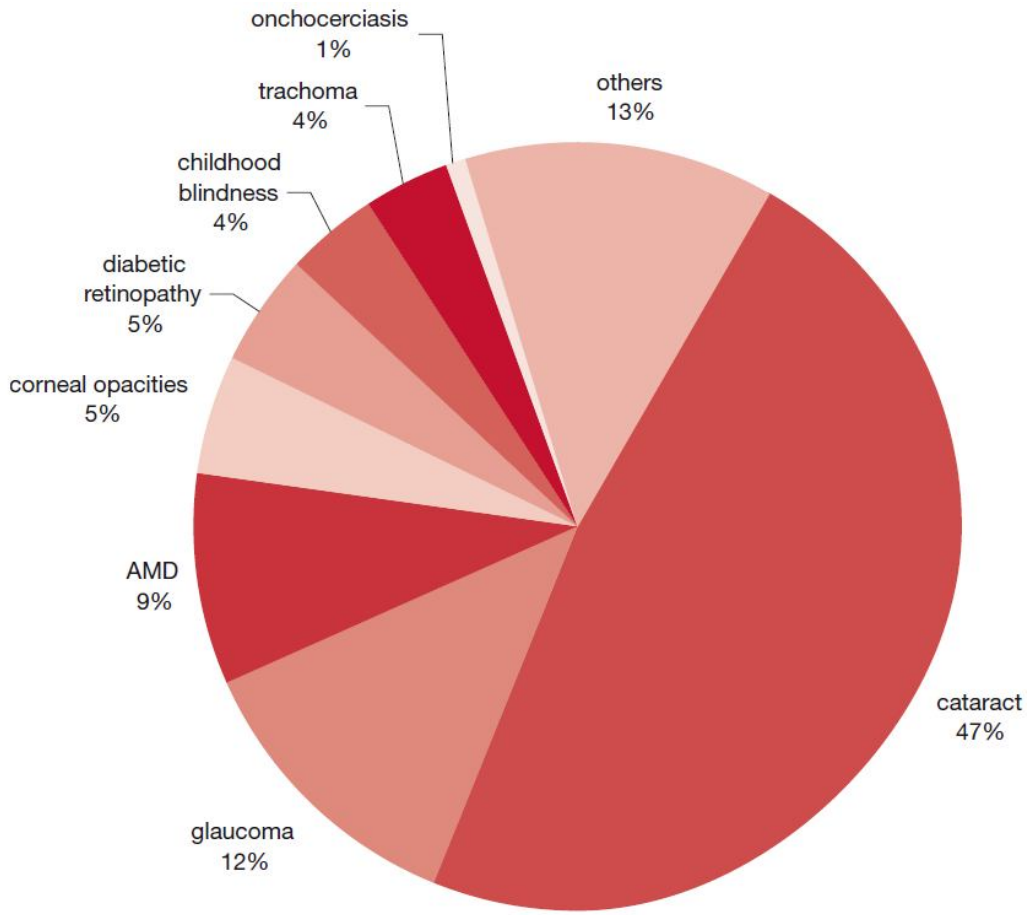

Figure 1.1 – Global causes of blindness due to eye disease; glaucoma is the second leading cause of blindness worldwide. The figure was reproduced from http://www.who.int/whr2001/2001/archives/2000/en/pdf/StatisticalAnnex.pdf accessed in June 2014.

There are various types of glaucoma (**Figure 1.2**), generally classified according to features of the disease, although increased intraocular pressure (IOP) is a common characteristic in most types. Glaucoma is commonly referred to as either primary or secondary; this refers to naturally occurring disease and disease caused by or through treating another existing condition. Primary glaucoma is further sub-classified through checking the angle formed between the iris and cornea of the eye using a method called gonioscopy (National Institute for Health and Clinical Excellence 2009). Angle closure glaucoma (ACG) is associated with a narrow angle between the iris and the cornea. When the angle is closed, the iris can block the trabecular meshwork, blocking the drainage canals and causing a build-up of fluid inside the eye, increasing the IOP (**Figure 1.3**). The onset can be quick and painful,

as in acute ACG, or it can be slow and asymptomatic as chronic ACG usually is. However, the most common form of glaucoma is when the angle to be wide, as in primary open angle glaucoma (POAG), which is thought to arise due to impeded aqueous outflow through the trabecular meshwork, causing a rise in IOP. Open angle glaucoma is an incurable progressive condition that requires continual monitoring after diagnosis and is managed through various medical interventions such as eye-drops or surgery to reduce the IOP. The progression of POAG tends to be gradual and symptomless in its early stages. Though most glaucoma is associated with increased IOP, glaucoma can occur when pressures inside the eye are at population normal levels (normal tension glaucoma; NTG). Another primary glaucoma is congenital glaucoma, a life-long condition contracted from birth. However, as this form of glaucoma is relatively rare, congenital glaucoma will not be a focus of the thesis. Common varieties of less prevalent secondary glaucoma include pseudoexfoliative and pigmentary glaucoma. Pseudoexfoliative glaucoma can be seen in eyes with pseudoexfoliation syndrome, which is characterised by deposition of microscopic granular protein fibers (which are like dandruff) in the anterior segment (the area between the cornea and iris) of the eye. Glaucoma occurs in this case where this pseudoexfoliative material blocks the drainage canals. Pigmentary glaucoma can occur in individuals with pigment dispersion syndrome in which pigment is shed from the back of the iris, which can then block drainage canals for ocular aqueous humour.

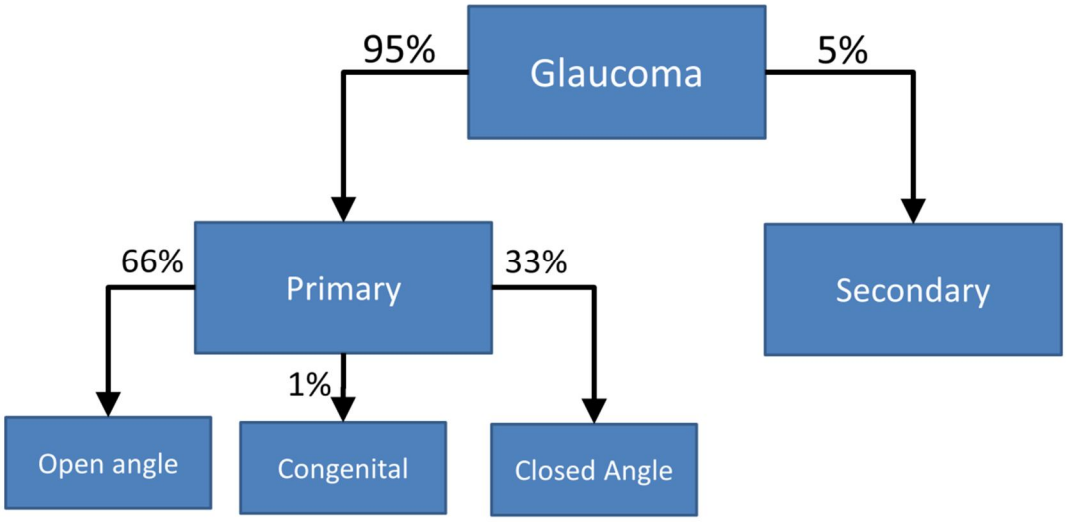

Figure 1.2 – A flow-chart showing the prevalence of different classifications of glaucoma. Pseudoexfoliative and pigmentary glaucoma are classed as secondary glaucomas. Diagram recreated based on figure 7.1 from Henson 2000

Regardless of the mechanisms behind the condition, and much remains unknown about the cause of glaucomatous symptoms, the product of the condition is invariably the death of ganglion cell axons in the retina at the back of the eye. With the death of these cells, the signal from photosensitive rod and cone cells on the back of the eye cannot be transmitted along the optic nerve to the brain. This results in patches in the field of view where vision is impaired otherwise known as scotomas. The rate at which vision is lost as a result is known as the rate of progression of the disease. Due to the fact that the brain tends to fill in information based upon the surrounding stimuli, patients often do not notice their scotomata until later stages of disease (Crabb et al. 2013). Added to the fact that glaucoma is usually characterised by gradual progression (Heijl et al. 2012a), does not tend to affect central vision later stages of the disease and the fields from both eyes cover the same central area (binocular summation) patients often do not notice they are losing their vision until late in the disease (Shaw 2005). As a result, patients can be diagnosed with visual impairment that can seriously undermine their quality of life (Ang & Eke 2007, Kotecha et al. 2012a), hence the condition's nickname as "the silent thief of sight". The severity of this condition, as well as the

fact that there is not yet any cure for the blindness caused by it, makes finding ways of detecting and preventing the disease progression imperative.

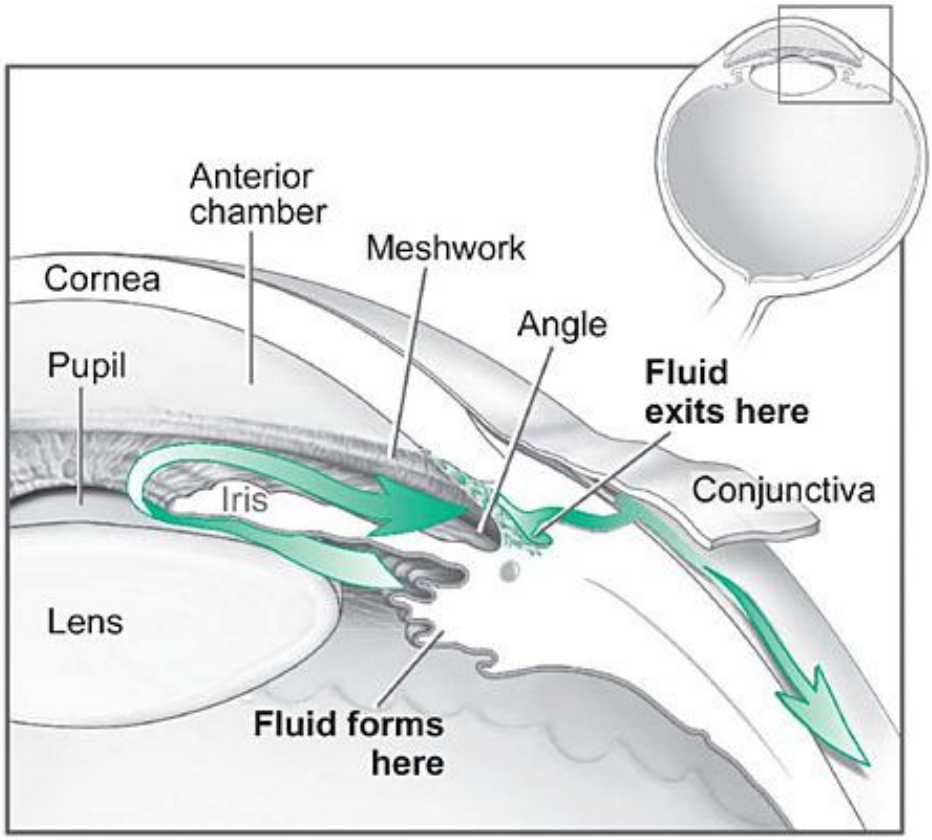

Figure 1.3 - The fluid pathway in the eye. Glaucoma is often related to the partial or complete blockage of the outflow of aqueous fluid through the trabecular meshwork (labelled "meshwork"). The "angle" refers to the angle between the iris and cornea. In open-angle glaucoma this is wide, whereas in angle-closure glaucoma this is narrow such that the iris presses against the cornea. Image taken from the National Eye Institute: http://www.nei.nih.gov/health/glaucoma/glaucoma_facts.asp accessed in June 2014.

1.1.1 Risk factors in glaucoma

There are various risk factors in glaucoma, yet the only modifiable one of these is an eye's IOP. Although not all individuals with high IOP have glaucoma (individuals with IOP over 21 mm Hg but no other signs of glaucoma are diagnosed as having ocular hypertension [OHT]) and not all glaucoma patients have high IOP (as is the

case in patients with NTG), it has been shown repeatedly to not only be a major risk factor for glaucomatous disease incidence and progression (Sommer et al. 1991, Gordon et al. 2002, Chauhan et al. 2008a, Kim et al. 2011, Jiang et al. 2012), but also a factor in the rate of progression of the disease (Heijl et al. 2012a, Medeiros et al. 2012, Chauhan et al. 2014). Large fluctuations in IOP have also been demonstrated to be a risk factor for glaucoma (Asrani et al. 2000, De Moraes et al. 2011a). How pressure relates to the death of ganglion cells is not fully understood, but its effects, along with the fact it can be changed, makes it an essential risk factor to consider.

There are various other factors associated with glaucoma that can be used to help identify at risk groups of the disease. Old age is very consistently linked with developing glaucoma, with over 50s a particularly at risk group (Mukesh et al. 2002, Gordon et al. 2002, Leske et al. 2007, Chauhan et al. 2008a, National Eye Institute 2010). Similarly, other non-modifiable risk factors such as family history of glaucoma (Leske et al. 2008) and ethnicity (National Eye Institute 2010) are also often found to be risk factors in disease incidence, with individuals of African descent particularly vulnerable. For instance, while population studies estimate incidences of definite OAG at around 0.1% per year in largely Caucasian Europe (de Voogd et al. 2005) and Australia (Mukesh et al. 2002) populations, these estimates were as high as 0.6% in the predominantly Afro-Caribbean Barbados Eye Study (Leske et al. 2001). Features of the eye such as exfoliation syndrome (Heijl et al. 2009) are also linked to development of glaucoma. Finally, there are various structural features of the eye that have often been linked to glaucoma, such as thin central corneal thickness (CCT) (Gordon et al. 2002, Medeiros et al. 2003, European Glaucoma Prevention Study Group 2007, De Moraes et al. 2011a, Medieros et al. 2012) and large axial length (which causes myopia or short-sightedness) (Jiang et al. 2012, Marcus et al. 2011) though the latter factor has not consistently been shown to be a reliable risk factor (Sohn et al. 2010). Interestingly, despite studies showing that individuals with OHT tend to have high CCT compared to those with glaucoma (Sobottka Ventura et al. 2001), attempts to correct IOP for CCT in prediction models have not yet been found to have great utility (Brandt et al. 2012). One possible

explanation is down to the fact that thicker corneas tend to give higher IOP readings using tonometry than individuals with thinner corneas.

There are various other suspected factors that are not universally acknowledged as risk factors for disease. For instance, increased systolic blood pressure has been linked to disease incidence (Jiang et al. 2012) whilst other morbidities such as diabetes (Mitchell et al. 1997, Gordon et al. 2002), vasospastic disease (Broadway & Drance 1998) and even sleep apnoea (Mojon et al. 2000) have all previously been linked, although their relationships with the disease is by no means proven (Chauhan et al. 2008a, European Glaucoma Prevention Study Group 2007). Ocular perfusion pressure (OPP) is another clinical characteristic that has previously been linked to developing glaucoma (Leske et al. 2008, Zheng et al. 2010), largely in an attempt to explain why not every glaucoma patient has high IOP. However, ocular blood flow is difficult to measure and has largely been estimated using a function of the blood pressure and IOP. Unfortunately, this measurement does not give any more information than blood pressure where IOP is also being accounted for (Khawaja et al. 2013). As a result, the status of OPP as an independent risk factor is still questionable.

1.1.2 Diagnosis of glaucoma

Glaucoma is diagnosed through using a battery of different tests: National Institute for Health and Clinical Excellence (NICE) Guidelines stipulate that all people suspected of having POAG or who even have OHT should have Goldmann applanation tonometry performed, CCT measured using pachymetry, gonioscopy performed, VF measurements using standard automated perimetry (SAP) and optic nerve head (ONH) assessment using stereoscopic slit lamp biomicroscopy (National Institute for Health and Clinical Excellence 2009).

Tonometry is the method used in measuring the main risk-factor of glaucoma, the IOP, so is understandably important in determining pressure targets for treatment. Though not necessarily directly related to glaucoma itself, in conjunction with high IOP, thin CCT can be linked to greater likelihood and speed of progression in glaucoma and is therefore required to be taken into account when determining

treatment severity (National Institute for Health and Clinical Excellence 2009). The ONH assessment allows clinicians to look at the health of the optic nerve, but SAP is the only means of directly investigating the actual visual health in terms of its impact on the patient. Due to its importance in assessing disease status and its relevance to the patient, it is measurements from SAP that will be predominantly looked at over the course of this thesis.

1.1.3 Treatment for glaucoma

Whilst, these risk factors are all related to glaucoma and could be useful for consideration in screening purposes, controlling the IOP remains the only means of treating the disease. Various clinical trials, including the Ocular Hypertension Treatment study (OHTS) (Gordon et al. 2002), the Early Manifest Glaucoma Trial (EMGT) (Heijl et al. 2002, Leske et al. 2003) and the Canadian Glaucoma Study (CGS) (Chauhan et al. 2010) have all demonstrated the utility of lowering the IOP to reduce the incidence and progression of glaucoma.

Management of IOP is performed either through medication (normally eye drops), laser treatment or surgery. Eye drops are certainly preferable, but can have various side-effects ranging from eye irritation to nausea. Non-adherence to treatment is a large issue in glaucoma (Gurwitz et al. 1993, Shaw 2005) for various reasons, including the fact that patients do not realise their vision is getting worse and therefore the importance of adherence. In addition, treatments can also be difficult to administer, particularly for elderly patients, who are unfortunately also the main demographic affected by glaucoma.

It is important to ensure that patient vision does not deteriorate to blindness or even visual disability within their remaining lifetime, so it is important not to under-treat patients. However, there is a significant risk of overtreatment in glaucoma too, which wastes limited clinical resources; a particularly pertinent issue given that the growing numbers of patients will increasingly stretch national health resources (Tuulonen 2013) and that the condition may not necessarily lead to visual impairment (Heijl et al. 2009). In addition, although patients would rather have

surgery than lose vision (Bhargava et al. 2006), it must be taken into consideration that treatment tends to become more unpleasant as severity increases. For instance, trabeculectomy can potentially result in increased risk of cataract, infection, blurred vision, bleeding, sudden, permanent loss of central vision and even secondary glaucoma if fluid drainage is prevented by scarring.

1.2 Monitoring glaucomatous vision loss

It is clearly highly important to monitor patient visual deterioration over time in order to evaluate whether or not treatment is required or needs to be escalated (Heijl 2013). The next section will briefly look at ways of doing so in clinical practice.

1.2.1 Structural measurements

One method of monitoring loss is to measure the changing structural characteristics of the eye as they change with disease. For example, examination of the optic disc and measurement of retinal nerve fibre layer thickness can be used. New, high-resolution imaging instruments such as Spectral-Domain Optical Coherence Tomography are starting to emerge from research laboratories in the hope of more accurately identifying disease progression (**Figure 1.4**). The ultimate advantage of structural methods is that they are objective, as they are not reliant on patient response. However, current imaging devices, though useful, are not a replacement for functional measurements in measuring glaucomatous progression (Strouthidis & Garway-Heath 2008, Gabriele et al. 2011). The largest issue is that it is difficult to reconcile structural measures to clinical outcomes that are meaningful to the patient, as they do not portray what the patient can actually see and therefore the impact of the disease. Moreover, there can be large discrepancies between structural and functional measurements of progression in glaucoma to the point of being largely independent of one another (Artes & Chauhan 2005). In addition, despite appearing precise, structural measurements are still subject to variability (Owen et al. 2006). Structural measurements are not presently acceptable for the evaluation of medical products for the treatment of glaucoma (Weinreb & Kaufman

2011). They are, nonetheless, useful tools in glaucoma diagnosis and monitoring when used alongside more functional means of measuring glaucomatous loss. However, this thesis will not focus on the use of these measurements.

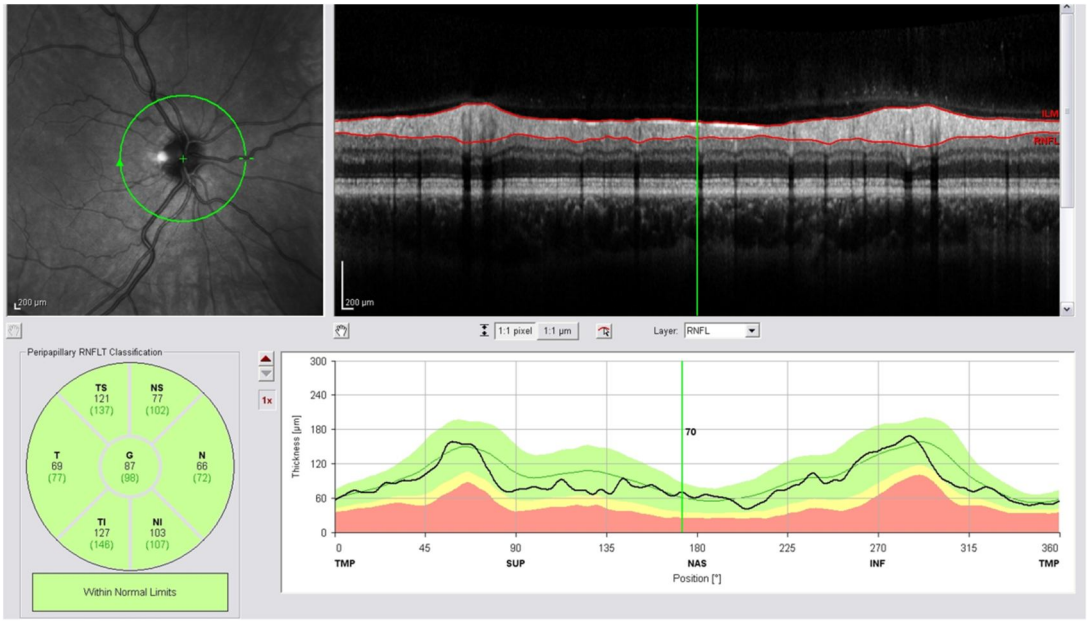

Figure 1.4 – Output from an OCT showing the Retinal Nerve Fibre Layer of my own right eye.

1.2.2 Perimetry

Measuring the functional progression of the glaucoma (that is, what the patient can actually see) is important in diagnosing glaucoma and monitoring its progression. Perimetry is the means by which the VF; of a patient is mapped (Henson 2000), and the only means of measuring functional progression. Other than reduction in IOP (which is the basis by which most new therapies are evaluated), VF measurements are the only accepted endpoints in the evaluation of new treatments for glaucoma; the Advanced Glaucoma Intervention Study (AGIS), Collaborative Initial Glaucoma Treatment Study (CIGTS) and EMGT being examples of major trials in which VF progression was the chief endpoint.

Measuring an individual's VF is simple in principle; the subject must first fix their eyes on a particular point and light stimuli of varying intensities are then displayed around the individual's field of vision. The subject must then communicate to the examiner whether they can see the light or not. Traditionally, manual methods of

doing this were performed including the Goldmann perimeter, but these approaches have largely been superseded by automated perimetry due to the fact that Goldmann perimetry is highly dependent on the examiner in terms of accuracy and bias. Given Goldman perimetry involves a moving stimulus, patients undertaking this test tend to lose fixation to a greater extent than those using automated perimetry where the pseudo-randomised location of the stimulus is less predictable (Heijl & Krakau 1977). Thus, static automated perimetry is a clinical gold standard in clinical practice as the more reliable and more reproducible option (Fankhauser et al. 1977), although Goldmann perimetry can sometimes still be used in cases where patients struggle to interface with automated perimetry.

1.3 Standard Automated Perimetry

There are various automated perimetric methods, but of these standard automated perimetry (SAP) is most commonly used in clinical practice and regarded as the clinical standard. SAP uses white lights as stimuli presented on a white background. Tests typically use static stimuli, which are flashed sequentially in a pre-defined grid of locations. The distribution of test locations can vary, but the two common VF testing patterns are at 6-degree even intervals in a 30-2 pattern (with 38 points in each hemisphere spanning the central 30 degrees) or a 24-2 pattern (27 points in each hemisphere spanning the central 24 degrees). The duration and size of stimulus is also fixed in static perimetry (commonly at 0.2 seconds and 0.43˚ of visual angle in diameter [Goldmann Size III] respectively). However, though SAP has been available from the 1970s, there have nonetheless been a number of other automated perimetry methods that have been developed, including Short Wavelength Automated Perimetry (SWAP), Pulsar (or Flicker Perimetry) and Frequency Doubling Perimetry (FDP or FDT) (Giangiacomo et al. 2006, Turalba & Grosskreutz 2010). However, research is still ongoing on the utility of these newer techniques so, for the moment, SAP remains the primary method for detecting VF progression and is hence the subject of this research. As with structural methods, alternative functional methods may possibly be helpful alongside rather than instead of SAP (Chauhan et al. 2008b).

There are three main machines used for SAP in the UK at present: the Humphrey Visual Field Analyzer (HFA; Carl Zeiss Meditec, Dublin, CA), the Octopus (Haag-Streit, Köniz, Switzerland) and the Henson (Elektron technology, Cambridge, UK). These three devices produce slightly different outputs, but generally perform the same function, to similar standards. The HFA is most commonly used in many large clinical centres in the UK, especially in a tertiary or referral setting where the goal is to monitor VFs in patients with glaucoma or who are at risk of developing glaucoma.

1.3.1 Measuring the Visual Field using Standard Automated Perimetry

SAP derives an estimate of the retinal sensitivity at various equally-spaced point locations in a patient's VF. There are two common types of testing: threshold and supra-threshold testing. In case-finding, supra-threshold testing is often used because it is quick; this mode of testing uses one or more light intensities and simply tests whether these stimuli can be seen in each test location. However, in glaucoma monitoring, where measuring changes in severity at each location is important, threshold testing is more commonly used. The aim of threshold testing is to try and find, for each tested location, the lowest level of light that it is possible for a patient to identify. In order to determine the thresholds of each test location, it is important to test at each location repeatedly.

Figure 1.5 – A colleague performing Standard Automated Perimetry on an Octopus Visual Field device

In SAP, the subject fixates on a central stimulus and indicates whether or not they are able to see other lights flashed in randomised locations around their gaze at varying intensities by pressing a button (**Figure 1.5**). The result is a series of measurements at each tested location called sensitivities (or thresholds), measured in decibels, an inverse measure of the strength of the stimulus (measured in candelas per metres^2). The VF threshold measurement ranges from 50 to 0dB with 0 dB representing perimetric blindness in that particular point of the eye (around 30dB is usually considered healthy). Generally, the results will be output in the form of a grid of numbers and a greyscale representing what parts of the VF are missing in black (**Figure 1.6**).

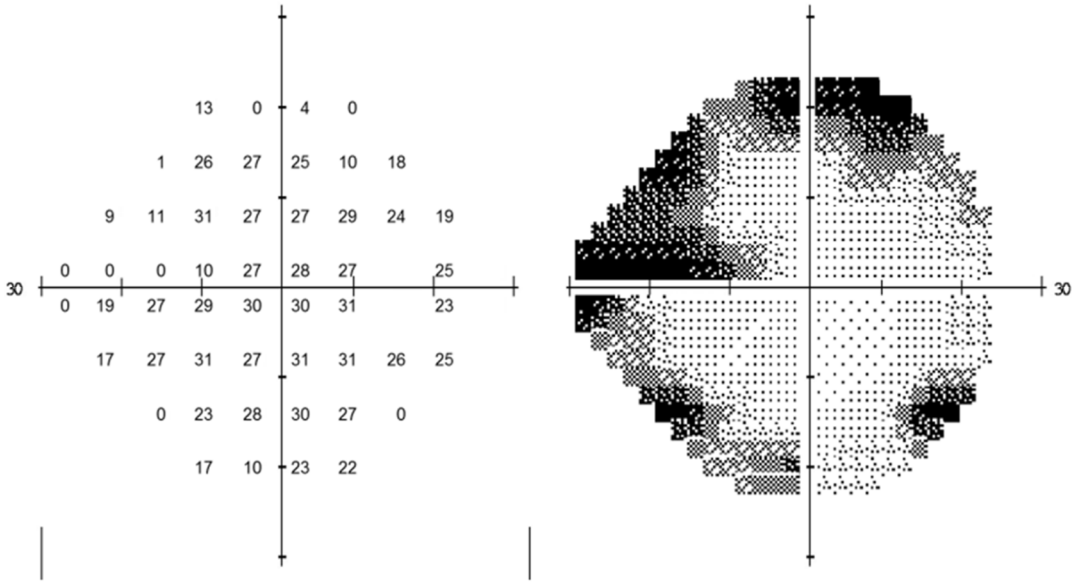

Figure 1.6 – Output from a Visual Field (VF) examination from a patient's right eye. The left grid shows the measured threshold sensitivities at 52 locations (excluding the blind spot), whereas the right grid is a greyscale with darker areas representing less sensitive parts of the VF.

In theory, patients will always see stimuli brighter than a test location's threshold, but fail to see a stimulus dimmer than a location's threshold. However, in reality, when a light brightness is presented at a level close to an individual's true threshold, there is a chance the patient may not respond as expected. In other words, there is a probability of a patient failing to register the stimulus even if it is visible to detect. The probability of a patient seeing or not seeing a given stimulus is therefore sometimes referred to as the frequency of seeing. The aim of testing is to try and find a light threshold at which a stimulus is seen 50% of the time.

1.3.2 Perimetric Testing algorithms

There are different methods for establishing VF sensitivities at each location. The earliest to be widely used in perimetry was full-threshold testing. In this method, an initial stimulus is presented at a test location. If a patient sees this presentation, the stimulus brightness is decreased by 4dB, whereas if it cannot be seen then the brightness is increased by the same amount. This continues until the status of whether the point can or cannot be seen changes (the first reversal). The stimulus

intensity is then decreased or increased back in steps of 2dB until the patient response changes once more (the second reversal). An average of the final and penultimate test sensitivities is recorded as the threshold for that test location. The order of the VF locations tested is randomised automatically in order to aid fixation (Heijl & Krakau 1975). Ultimately, this method is thorough, but it can also take a relatively long time to complete (about 10 to 15 minutes for each eye for a 24-2 test pattern) (Bengtsson & Heijl 1998a, Bengtsson & Heijl 1998b, Nordmann et al. 1998, Wild et al. 1999) during which it can be difficult for the patient to maintain their attention. As a result, it is not suitable for all patients.

Though, in theory, the increased amount of testing is supposed to improve test precision, some researchers speculate that the resulting fatigue from the length of a full threshold testing could exaggerate defects or even result in the finding of defects that do not exist (Bengtsson & Heijl 1998b, Turpin & McKendrick 2011). Thus, faster methods have been devised in order to reduce test time. The most successful and most widely adopted of these techniques is known as the Swedish Interactive Thresholding Algorithm (or SITA as it is commonly known) for the HFA. This Bayesian technique relies on a similar principle to the full threshold methods, but seeks to cut out the unnecessary testing time by reducing the number of presentations (Bengtsson et al. 1997a).

SITA Standard begins with prior information about the sensitivity of each location before the test starts, using age-corrected normal values, anticipated frequency-of-seeing curves and correlations between sensitivities at each test location (Bengtsson et al. 1997a). However, the test begins presenting stimuli in the same way as full threshold testing, measuring the first four “primary” points (or seed points) in a stepwise manner; one in each quadrant of the VF 9 degrees away from fixation in both axes. As with full threshold testing, after presenting an initial stimulus the test simply seeks the threshold sensitivity of each test location in the eye through increasing or decreasing the light sensitivity presented in steps of 4dB until the patient response changes. Once the first four points are estimated, the sensitivities of the other test locations can be estimated and these estimations are amended as the testing proceeds. In other words, the threshold of one point acts

to “predict” the probability of seeing a threshold for the others, so each test is actively changing the predicted threshold of the other points. A second reversal only occurs if the difference between adjacent point thresholds is larger than a pre-calculated value based on the patient’s demographic known as the error related factor (ERF) (Bengtsson et al. 1997a). The exact details of how the ERF is calculated have not been disclosed, but it is calculated automatically by most perimetry machines. During the test, response times are continually recorded and used to adjust the duration of test presentations. At the end of the test, thresholds may be adjusted slightly in post-processing according to thresholds of adjacent test locations and changes in reaction times during the test.

The large advantage of the SITA Standard is that it can halve the duration of testing for many patients (Bengtsson & Heijl 1998a, Nordmann et al. 1998, Wild et al. 1999). However, there is a consistent discrepancy between the SITA Standard and full threshold methods in that the former algorithm is consistently more optimistic in terms of threshold measurement by around 1dB (**Figure 1.7**) (Artes et al. 2002, Bengtsson & Heijl 1999). As a result of the time saved using it and the fact that it has similar repeatability to full threshold testing (Artes et al. 2002, Bengtsson & Heijl 1999), SITA Standard is now the most commonly used threshold detection method.

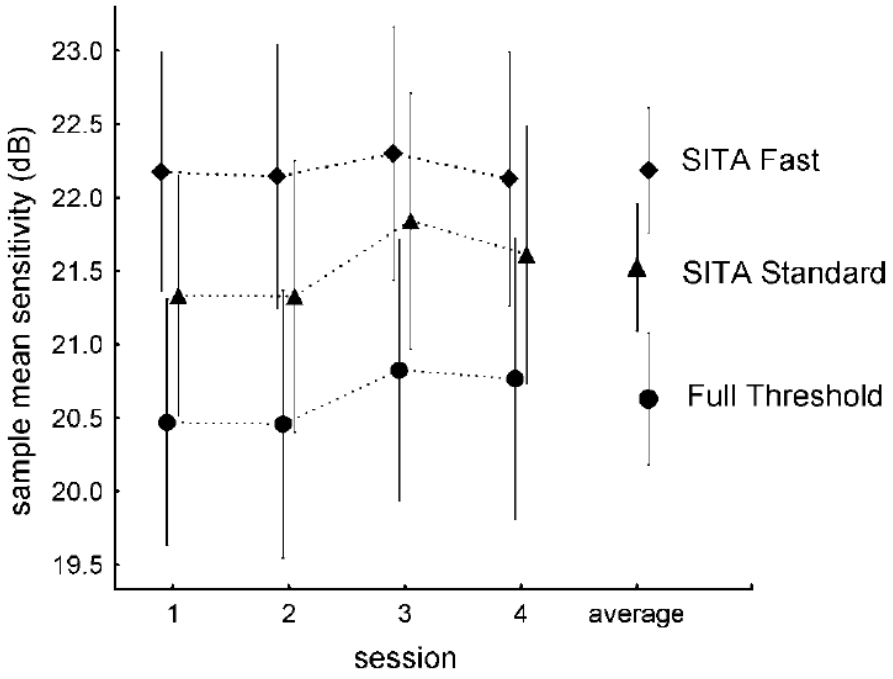

Figure 1.7 – The mean sensitivities for Full threshold, SITA Standard and SITA Fast tests taken on the same patients in a study by Artes et al. 2002. SITA Fast and Standard tests tended to yield higher, more optimistic thresholds than full threshold testing. Image taken from Artes et al. 2002.

Other even faster methods have been devised by researchers in order to enhance the speed of testing further. SITA Fast, in particular, is commonly reported to reduce testing times to below 5 minutes (Bengtsson & Heijl 1998b, Nordmann et al. 1998, Wild et al. 1999, Pierre-Filho et al. 2006), which could potentially have clinical utility in terms of saving time in clinical practice. SITA Fast was designed to be equivalent in terms of accuracy to Fast-threshold strategies, such as Fastpac (Glass et al. 1995), which uses 3dB steps and one reversal instead of two reversals in full threshold testing, but it is significantly faster (Bengtsson & Heijl 1998b). However, the test algorithm itself is similar to SITA Standard, the main difference being fewer reversals. Only one reversal is used (as opposed to two) unless the difference between the estimated and expected thresholds is greater than more than 12dB. In addition, the algorithm can terminate testing of any test location even earlier provided there is at least one positive response at a test location and the measurement error is below the ERF (Bengtsson & Heijl 1998b). Unsurprisingly given the nature of its design, there is a suspicion that SITA Fast may have less

sensitivity in identifying VF defects than full threshold testing and SITA Standard and there is some evidence that the results from this test are less repeatable (Artes et al. 2002) (**Figure 1.8**). It is uncertain whether the potential for increased testing and the impact on fatigue compensate for the natural reduction in precision as a result of less testing.

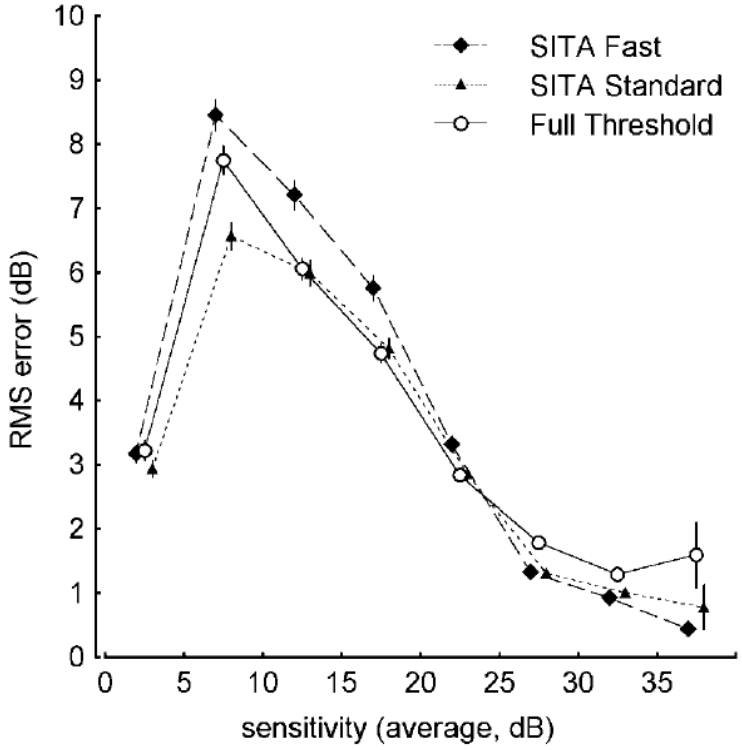

Figure 1.8 – The test-retest variability about full-threshold measured thresholds using Full threshold, SITA Standard and SITA Fast VF testing. The variability is greater for SITA Fast than Full threshold, but SITA Standard performs relatively well by comparison. This image is taken from Artes et al. 2002.

1.3.3 Reliability Indices

In an ideal world, perimetry should produce accurate results on every occasion, but results that are not reflective of a patient's actual VF can be produced, due to loss of attentiveness, inexperience, overenthusiasm or not fixating on the central point well enough. It is therefore important to evaluate the reliability of the VFs measured before using them to inform clinical decision making.

Every testing algorithm described above have methods of evaluating the reliability of the test itself through looking for false negatives (FN), false positives (FP) and

fixation losses (FL). False positives are a measure of how 'trigger-happy' the patient is, measuring how likely they are to indicate observation of a stimulus without seeing it. In full threshold testing they are tested through simply having occasions where the device pretends to change the stimulus location, but shows no light. To cut testing times, the SITA algorithms do not use catch trials, but judge false positives by the reaction speed of the patients, such as when a subject responds before they have had time to see and react to the stimulus. False negatives are a measure of inattentiveness of the patient – the devices show a light in an area where testing has already confirmed the threshold and give a brightness that the patient should be able to detect. Fixation losses meanwhile test how accurately the subject is fixating at the fixation target by presenting stimuli at the location of their physiological blindspot; this is the area corresponding to their optic disc, which has no photoreceptors.

The major clinical trials, such as AGIS, CIGTS and EMGT have used all these measures to assess field reliability (The Advanced Glaucoma Intervention Study Investigators 1994, Gillespie et al. 2003, Heijl et al. 2008), yet the criteria applied to them remain arbitrary and vary between trials. For instance, for the AGIS and CIGTS trials, a scoring system was used which meant that patients could theoretically pass with false positive or negative rates in excess of 33% (The Advanced Glaucoma Intervention Study Investigators 1994, Gillespie et al. 2003), whilst EMGT trials did not look at the FN rate at all (Heijl, Bengtsson et al. 2008). In addition, other reliability criteria such as the total number of questions asked (The Advanced Glaucoma Intervention Study Investigators 1994) (longer tests imply greater uncertainty in measuring the threshold) and short term fluctuation values have been used (The Advanced Glaucoma Intervention Study Investigators 1994, Gillespie et al. 2003).

There is evidence to suggest that none of the reliability indices provide an accurate insight into a patient's performance. For instance, Bengtsson found that none of the main three metrics for reliability contributed significantly more information to test reproducibility than the level of VF loss itself (see Section 1.3.4). The only metric that linked to test reproducibility at all were FNs (Bengtsson 2000).

However, there were indications in the same study that this relationship was most likely to be due to a significant association between FNs and VF loss (Bengtsson 2000). Similarly, Shao et al. found FNs to be the best index for predicting test reproducibility accounting for severity of VF loss, but also found that none of the reliability indices are good predictors of overall test-retest variability (Shao et al. 2011). Montolio et al. meanwhile, found false positives to be the most crucial index, with each 10% increase in FPs being estimated to increase threshold estimates by 1dB (Junoy Montolio et al. 2012). Fixation losses have been shown to contribute to test variability without being a major contributor (Henson et al. 1996, Junoy Montolio et al. 2012).

Overall, there is likely no single "best" reliability index in assessing progression; if a VF is unreliable in any way then this has the potential of hindering the ability to detect progression, whether it gives the impression that the patient's VF is better or worse than in reality, thereby leading to impaired clinical judgement. For this reason, clinicians should be aware reliability indices when taking VFs. Additionally, the instructions given to the patient, the correction of spherical ammetropia and patient attention may also have a significant bearing on the result and these indices should not be relied upon exclusively (Chauhan et al. 2008b).

1.3.4. Problems in monitoring Visual Field deterioration in perimetry

The output produced from SAP can be confusing as it contains a huge amount of data and information, and it is not always particularly obvious how large changes are from one VF assessment to the next. In spite of all the efforts in perimetry to accurately detect the threshold of a patient accurately, there is still a lot of "noise" in the measurements of thresholds, which is what makes measuring the rate of progression non-trivial, particularly in areas of the VF where vision is worse and this variability is greater (Henson et al. 2000, Artes et al. 2002, Russell et al. 2012a). As a result, there exists test re-test variability between VF tests that needs to be taken into account when analysing results.

Measuring a threshold where a patient can only see 50% of presentations can often be hard to find explicitly. Using a technique called method of constant stimuli to measure stimuli accurately (Laming & Laming 1992), Gardiner and colleagues found that some locations with measured thresholds using SAP did not have a stimulus intensity associated with them that could be seen 50% of the time, which led them to conclude the probability of seeing a stimulus did not increase appreciably regardless of how bright the stimulus was at measured 'thresholds' of below 19dB (Gardiner et al. 2014). In other words, they concluded that the likelihood of response at 19dB or brighter is governed by chance rather than giving any information on the retinal sensitivity at that location. If this is true, then perhaps there is an argument for incorporating a new lower limit for sensitivities in VF testing (e.g. setting 20dB as the lowest possible measurement), which would perhaps reduce variability in the calculation of progression indices. In the studies included in this thesis, however, it is assumed that there is still some information to be gained from lower threshold values in VF testing.

Furthermore, the psychophysical nature of the tests means that learning effects need to be taken into consideration, as patients often improve in their ability to participate in perimetric testing with experience (Wild et al. 1989, Heijl et al. 1989, Heijl & Bengtsson 1996). As a result, measured thresholds tend to increase over time, sometimes persisting long after the first three tests (Wild et al. 1989). As the first test is often particularly deflated (one previous study reported an average increase of as much as 2.6dB in MD between the first two tests in perimetry naïve glaucoma patients [Heijl & Bengtsson 1996]), discounting the first VFs remains good practice when assessing glaucoma progression.

In spite of all the problems associated with judging progression, many clinicians nonetheless assess progression 'manually' using their experience through comparing SAP printouts. However, decision-making, even among expert clinicians, can be inconsistent (Viswanathan et al. 2003, Tanna et al. 2011) and concordance between clinicians has been shown to increase substantially when Progressor software is used (Viswanathan et al. 2003). As a result, it is clear that clinical decision-making in evaluating progression status can be improved with the

thorough understanding and utilisation of the available software and analysis methods to facilitate this task (**Figure 1.9**).

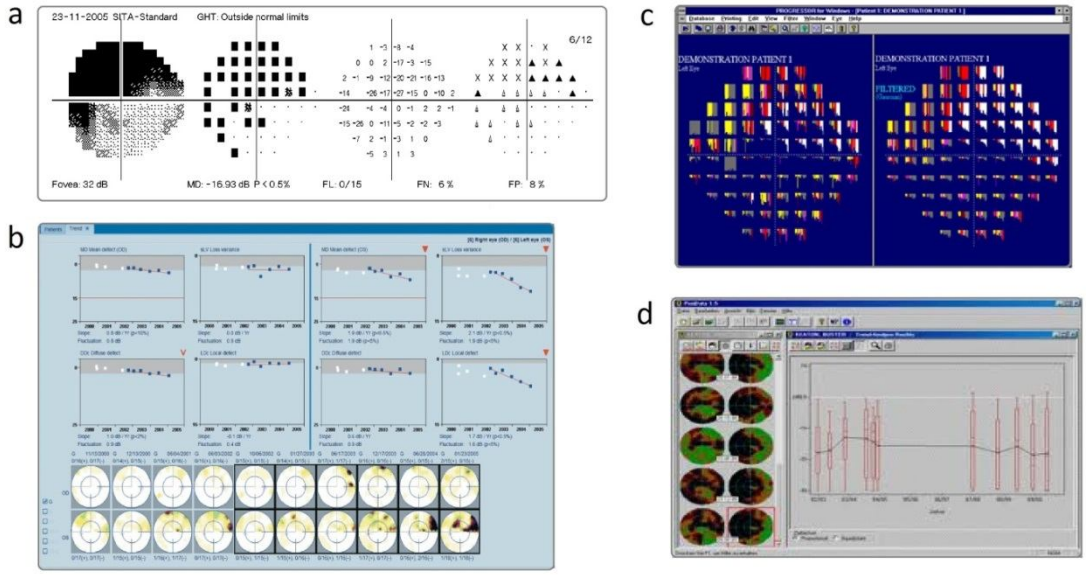

Figure 1.9 – A range of analytical tools available for use in clinical practice a) Statpac 2's Glaucoma Probability Analysis for the HFA, b) Eyesuite Analysis software for the Octopus perimeter, c) Progressor and d) Peridata's boxplot trend analysis

1.4. Global indices

Given the difficulty of thinking about 54 points simultaneously, it is often desirable to have a single global index to summarise the amount of vision loss in an eye. As a result, there are a range of global summary indices that have been developed for the purpose of measuring levels of VF damage.

1.4.1 The Mean Sensitivity, Mean Defect and Mean Deviation

The simplest way of summarising all of the VF sensitivities into one single index is to simply take their arithmetic mean. This is called the mean sensitivity (Flammer 1986). However, the disadvantage of using the mean sensitivity is that it is unclear whether a patient sensitivity is normal or not. Without a reference value, it is impossible to tell what is meant by a mean sensitivity of 26dB for a particular eye.

This is especially pertinent given the fact that the retinal sensitivity decreases naturally with age (Heijl et al. 1987).

As a result, every type of perimeter has a normative database of VF thresholds to compare measured VF thresholds against at a given location and for a given age. Measured thresholds are then compared with these average thresholds of the rest of the population and the differences between these for each location are known as total deviations (TD). Taking the arithmetic mean of these TD values gives the mean defect (Flammer 1986). However, it is well known that the variability in the periphery of the VF tends to be higher than in the centre around the point of fixation (Heijl et al. 1987) and the mean defect does not take into account these differences. In addition, it could be argued that the more central test locations are more important in the context of patient life and therefore require greater weighting.

The most commonly used and widely understood metric for summarising damage in an individual's VF is the Mean Deviation (MD) index (Artes et al. 2011). The MD is much like the mean defect except weighted to give more prominence to the less variable central test locations (Heijl et al. 1987). This allows for a more accurate summary of the eye's overall visual capabilities and it is therefore commonly utilised over the mean defect.

These statistics give a crude, yet understandable, single statistic to give an indication of how an individual's global VF sensitivity is compared with that of a "normal" person of that individual's age. However, the problem with using the MD (as well as the mean defect and sensitivity) and summary measures in general is that they are only useful in terms of summarising the defectiveness of the whole VF, when, in glaucoma, it is often the size and position of localised damage that are of most interest (**Figure 1.10**).

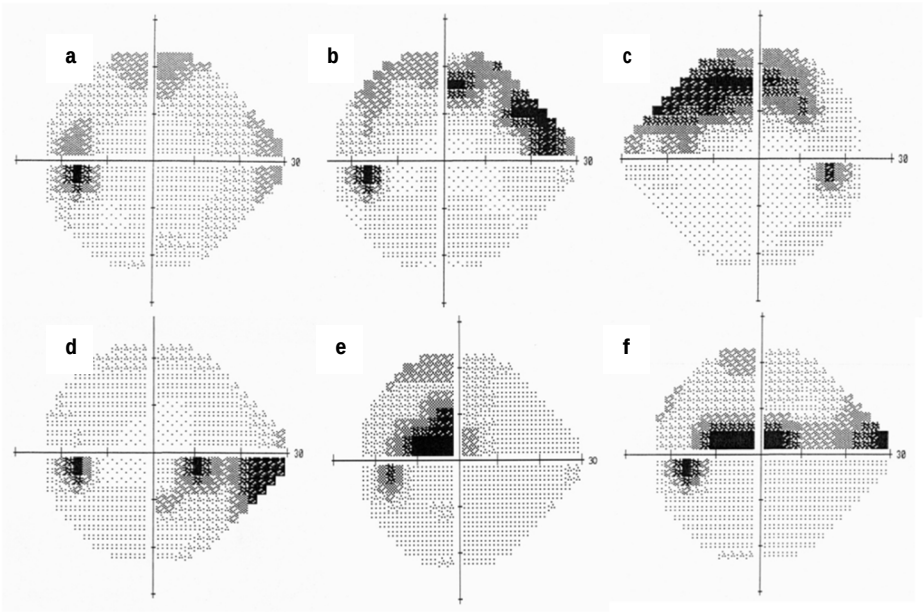

Figure 1.10 - HFA greyscale representations of six different eyes. Clearly all the VF defects are different and might impact on the patient's day-to-day function differently. Yet all of these VFs have the same MD value of -5dB. This illustrates the limitation of using a global index, or a single number like MD to summarise the VF, because all spatial information about the defect is lost. For instance, the visual function of a patient with a central defect (such as patients e and f) is likely to be more compromised than that of a patient whose VF is affected more peripherally (such as patients a to d).

1.4.2 Total Deviation Map

Humphrey print-outs have two TD maps which show where defects in the VF are located. One is a map of measured TDs at different locations in the VF. The second is a probability map where the measured TDs are compared to the distribution of thresholds in healthy eyes (the normative database), which are annotated according to the percentile below which they fall. Thresholds measured in the bottom 5% of the sample population flagged visually with different symbols for the sub 5, 2, 1 and 0.5 percentiles (**Figure 1.11**). This is an effective visual way of communicating the VF of one eye when used in conjunction with the summary statistics described previously.

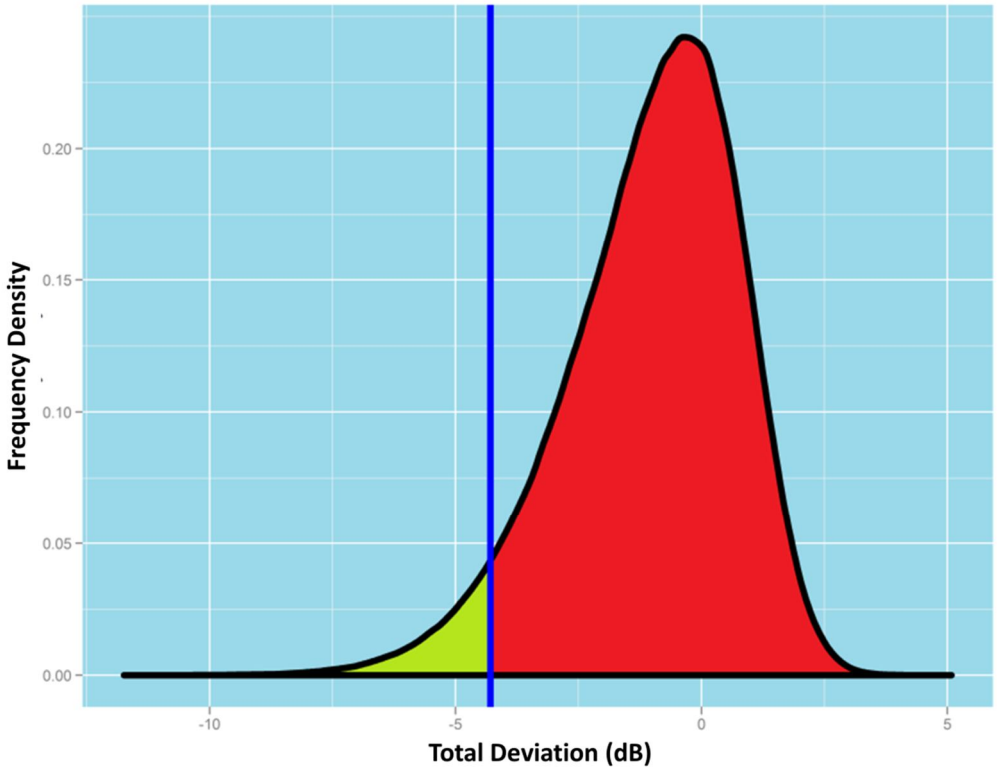

Figure 1.11 - The above distribution represents the variation in the total deviation value (TD) of a given point in a monocular VF. If a measured TD value falls within the green area (i.e. below the fifth percentile), then there is a significant probability of damage and the location will be flagged in the TD plot. This figure was previous published in the European Ophthalmic Review (Saunders et al. 2013).

1.4.3 The Pattern Deviation and Pattern Standard Deviation

There is a notable flaw with the methods of measuring VF loss discussed so far. Cataracts are a common eye condition in old age that causes a loss in sensitivity throughout the VF and using these methods exclusively it is hard to differentiate the effects of cataract and open-angle glaucoma. Additionally, a patient may simply have thresholds that are unusually low for someone of their age. As a result, the pattern deviation (PD) plot has also been devised and is commonly consulted by clinicians in order to assess whether defects are likely to be localised glaucomatous ones or a result of general VF worsening (**Figure 1.12**). To this end, it references each TD value against the 7th highest TD value and checks what

percentile the difference between these two values are. In other words, the PD attempts to compare focal VF loss with diffuse VF loss. As with the TD plot, any readings in the bottom 10% of the population are flagged.

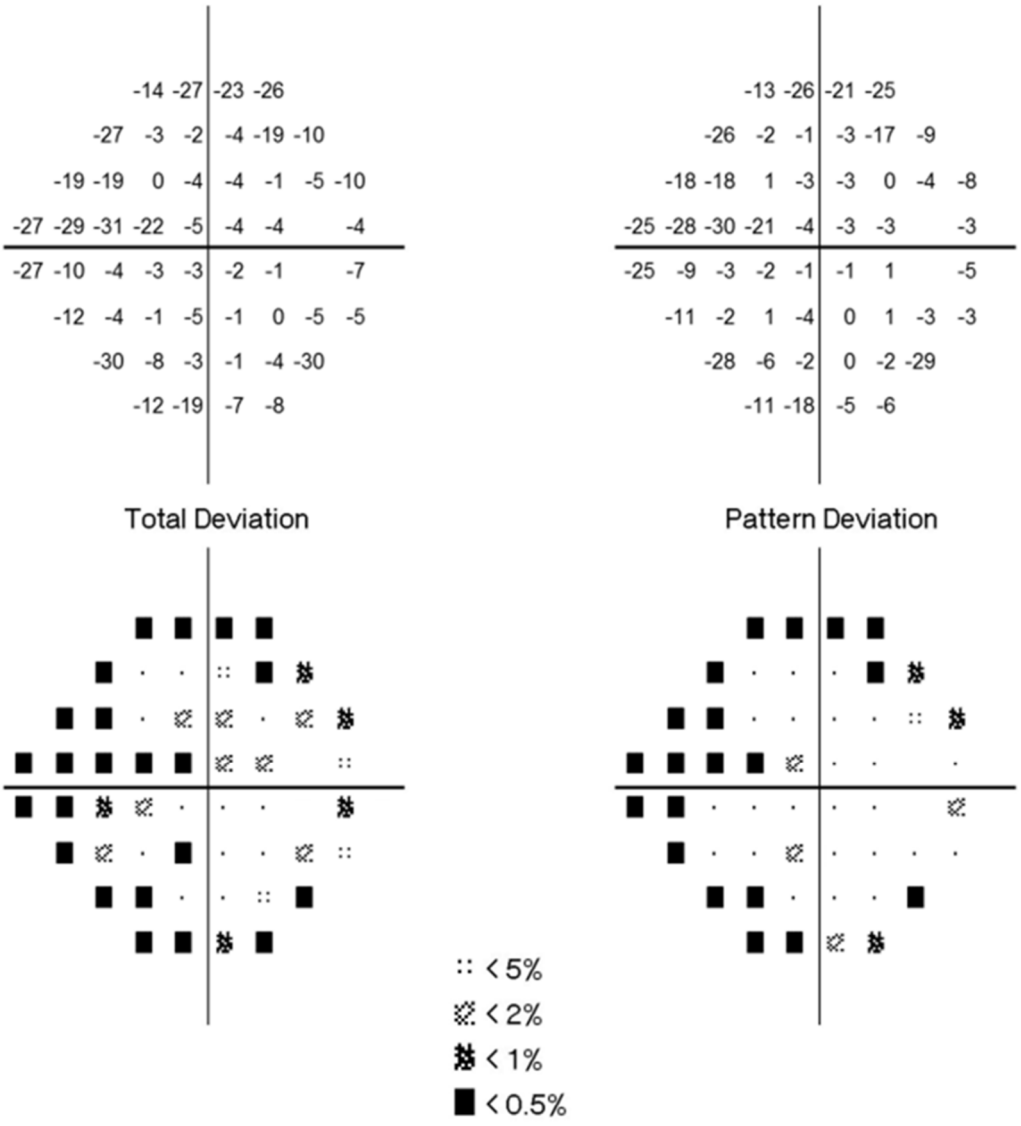

Figure 1.12 – Total and Pattern Deviation plots. The top grids show the number of decibels that each threshold deviates from the expected threshold, whereas the bottom grids show the percentiles of the estimated distribution of a normal population the thresholds reside below. Pattern deviation plots correct for diffuse loss in the eye that often results from cataract; a common co-morbidity in old age.

The pattern standard deviation (PSD) measures the amount of variability within the VF, meaning that it can be useful for differentiating glaucomatous focal loss from diffuse cataract loss. Specifically, it sums the absolute difference between the

sensitivity at each field location and the normal age-normal sensitivity corrected for the patient's MD. A high PSD thus implies that there is a great deal of variation between points measured in the VF, which is more indicative of typical glaucomatous field loss, whilst if each location is uniformly depressed (as in diffuse loss) the PSD is low. However, this measure is ineffective at the end-stage of glaucoma, because it will decrease as defects become more homogenous across the VF. Thus, the PSD not a good indicator of overall damage, though it can still be helpful when used alongside the MD (Brusini & Filacorda 2006).

1.4.4 The Visual Field Index

The Visual Field Index (VFI) is a relatively new summary measure that seeks to quantify glaucomatous damage, in later software upgrades of the HFA (Bengtsson & Heijl 2008). It is similar to the MD, though there are various notable differences. The VFI seeks to be user friendly to clinician and therefore attempts to estimate the percentage of vision the patient has left. Vision with no discernable defect is categorised at 100% with 0% signifying perimetric blindness. The VFI is weighted more towards the central VF than the MD, operating on the principle that the central part of the vision is of highest importance in terms of quality of life. Finally, the PD is utilised rather than the TD to calculate this statistic.

There are therefore two main benefits of using the VFI to monitor VF loss over other measures such as MD: firstly, it measures only damage that is related to localised VF defects, thus giving it relative immunity to the confounding effects of cataract, and, second, it prioritises the central VF more, therefore giving a better representation of how visual loss is likely to impact on visual function. Bengtsson and Heijl point out that using a percentage scale makes the result more relevant to patients than traditional decibel measurements (Bengtsson & Heijl 2008). However, Artes et al. have pointed out a number of flaws with the statistic (Artes et al. 2011). One key issue is that there is evidence that the VFI is overly-optimistic in estimating the proportion of vision left when compared with expert opinion, which

suggests that the statistic can be somewhat misleading in representing how well a patient can see (Artes et al. 2011). A possible reason behind this is that the reliance of the VFI statistic on pattern deviation is thought to fail to take into account the fact that glaucoma does also cause diffuse loss (Henson et al. 1999), which is ignored using this statistic. Thus, the VFI can risk underestimating the overall effect that glaucoma has on the eye (Artes et al. 2010) a fault also acknowledged by Bengtsson and Heijl (Bengtsson & Heijl 2008). Thus, though the VFI may potentially be useful, the MD is still a gold-standard in summarising VF loss into a single statistic.

1.4.5 Issues with global indices

Overall, though summary measures are useful in terms of having a singular measure representing how badly an individual's vision has been degraded by glaucoma, a universal problem for all of these measures is that they waste data and ignore important spatial information. Furthermore, it is difficult to tell how changes in these measures relate to visual disability. For example, what exactly does a drop in VFI from 100 percent to 97 percent mean? Thus, criteria must be devised in order to indicate whether glaucomatous progression is taking place at a dangerous rate or not. It is also accepted that summary measures are relatively insensitive to change when subject to analysis due to the fact that it averages the healthy parts of the VF as well as the unhealthy sectors to calculate the measure deterioration (Smith et al. 1996, Wild et al. 1993).

1.5 Event and Trend-based analyses

Whatever measure is utilised to summarise damage, it is important to be able to have some analysis method for establishing the difference between VF progression and differences in VF measurements due to variability. The analysis methods for doing this can generally be divided into two categories: event and trend-based analyses. There is often much debate between which type are better to use. Overall, there are clear differences between these two families of methods as a

result of how they treat measurements taken between the first and last VF. This gives them unique properties that have to be considered before evaluating which set of methods have the best utility in detecting glaucomatous progression.

Event based analysis involves taking a baseline and comparing every subsequent test against this reading. Any significant difference between the baseline and latest reading is considered to be due to progression. On the other hand, trend based analyses are an evolving process in which all VFs are analysed using linear regression to assess the rate and significance at which the measurement is changing.

1.5.1 Staging glaucoma patients

One means of marking progression is through using some staging system to mark when glaucomatous eyes have progressed beyond a certain level. Although useful for categorising glaucomatous eyes at diagnosis and for categorising damage in various studies, they are generally too insensitive to change to be practical in terms of monitoring progression. These methods are described in depth by Susanna Jr and Vessani (Susanna Jr. & Vessani 2009).

The most commonly used method of staging glaucoma is the Hodapp-Parrish-Anderson (H-P-A) index, which categorises defect severity into three stages (Early, Moderate and Severe) based upon MD, numbers of points below the 5% in the pattern deviation and health of points specifically in the central 5˚ (Hodapp et al. 1993). Though over 20 years old, this method of categorisation remains useful for roughly categorising VF loss and is still popular for use as a standard for VF defect severity (Elbozan Cumurcu et al. 2010, Labiris et al. 2010).

However, the fact that there are only three stages means that this method is not always helpful in the context of monitoring disease. In order to deal with this, some have adapted the H-P-A criteria, to contain more categories (Mills et al. 2006, Kobelt et al. 2006). Others have created newer more complex methods of categorising loss. For instance, Brusini & Filacorda devised the Glaucoma Staging

System (GSS), which incorporates MD and PSD to determine the stage of the defect (Brusini & Filacorda 2006). One very recent categorisation method is the University of Sao Paulo Glaucoma Visual Field Staging System (USP-GVFSS), which incorporates the VFI, the proximity of damage to fixation, the number of hemifields affected and whether the damage is connected to the blindspot into one string of code (Susanna Jr. & Vessani 2009). However, the increase in utility that has come in increasing the number of categories come at a cost of making the categorisation criteria more complex and as a result none of these methods have made a large impact in clinical practice.

1.5.2 Pointwise scoring criteria

Pointwise scoring criteria (examining each VF location separately) are sometimes used for determining progression though usually only in clinical trials, as they have good diagnostic specificity (probability of diagnosing healthy eyes as non-progressing) (The Advanced Glaucoma Intervention Study Investigators 1994, Heijl et al. 2008, Musch et al. 1999, Heijl et al. 2003, Ernest et al. 2011, Vesti et al. 2003). Two of the most famous criteria for detecting glaucoma progression were developed in order to analyse the results of two large scale clinical trials, AGIS and CIGTS, and hence are named after those trials.

However, both of the AGIS and CIGTS scoring methods, though specific in terms of their criteria of progression, both suffer from similar flaws to those of the summary methods described earlier. Specifically, these scoring systems can be affected and lowered by cataract, whilst there is also a fundamental loss of detail that occurs when summarising disease severity by using a single number in a disease that fundamentally affects the VF in a localised manner. As a result, in clinical practice, more sensitive event and trend based analyses tend to be utilised in order to estimate VF progression.

1.5.3 Glaucoma Change Probability Analysis

Perhaps the most useful form of event analysis to date is the Glaucoma Change

Probability method (GCP), which is sometimes called Glaucoma Probability Analysis or Guided Progression Analysis and is available from Statpac 2 GPA software for the HFA. This method takes the first three VFs and then averages the thresholds from the two most reliable VFs to attain a baseline reading for each test location. This baseline is compared with all subsequent VF tests and the difference between them is assessed using the GCP map, which designates the amount of variability from the average threshold one would expect from the baseline value, calculated from typical population variability. Any significant difference between the baseline and latest reading that is confirmed in two subsequent tests is considered to be due to progression and a new baseline is taken. Typically with this type of analysis, three points or more being shown as consistently defective in three tests are required to mark progression. An adapted version of the GCP map known as the pattern deviation GCP map has been used for the EMGT, which accounts for cataract based defects similarly to how the pattern deviation isolates localised effects from TD plots (Bengtsson et al. 1997b, Heijl et al. 2003).

The GCP method is one of the most sensitive methods for detecting glaucomatous progression (Vesti et al. 2003) and is in good agreement with expert opinion on progression status (Heijl et al. 2008). The test's specificity (its ability to correctly diagnose non-progressors) has been found to be slightly worse than less sensitive methods such as the CIGTS and AGIS methods, but FP rates are nonetheless good, with one study estimating FP rates of progression of just 2.6% (Artes et al. 2014). Furthermore, the GCP method has substantially higher sensitivity in detecting change than these scoring methods (Vesti et al. 2003, Heijl et al. 2008, Ernest et al. 2011). Importantly, given the context of wanting to diagnose VF progression as quickly as possible to prevent future impairment, GCP analysis is also much quicker in diagnosing progressive change than the CIGTS and AGIS scoring methods (Heijl et al. 2008), though only under the condition of frequent testing (Vesti et al. 2003). It is important to further note that the use of the GCP map is based upon the population variability; some patients are more consistent VF test takers than others (Heijl et al. 1987) and using a population based measure ignores this. As a result,

patients that are exhibiting signs of progression but are consistent test-takers may not get flagged as early as they should.

1.5.4 Trend-based analyses

Event analyses are useful in determining whether progressive change is occurring, but cannot be used to predict future outcomes, which may be important in informing treatment decisions. In addition, they only tend to look at two VF observations in determining whether progression has occurred and this seems wasteful of the large amount of data collected in long term monitoring. As a result, incorporating trend-based analyses that can be used to anticipate future VF status are popular.

The simplest type of trend analysis is to monitor how summary measures change linearly over time using ordinary least squares regression (OLSR) and this is easily accessible for clinicians using software such as Statpac 2 and EyeSuite Progression Analysis software for the Octopus (**Figure 1.9**). Recently, this approach has been specifically advocated with the VFI in particular (Bengtsson et al. 2009), but regression of summary measures has been occasionally criticised for being relatively insensitive to measuring progression (Smith et al. 1996, Wild et al. 1993). Furthermore, summarising all of the points in the VF will tend to result in equally weighting the damaged and undamaged parts, which will inevitably not be as sensitive as concentrating on the faster progressing sections of the field. However, this approach can be nonetheless useful in giving an impression of the future status of an individual's VF.

A more sensitive commonly utilised approach is Pointwise Linear Regression (PLR), available from PROGRESSOR (Progressor Medisoft, London, UK), EyeSuite (Haag-Streit, Köniz, Switzerland) or Peridata (Peridata Software GmbH, Huerth, Germany) software (**Figure 1.9**), which detects the rate of progression of each point in the VF (Fitzke et al. 1996). As the name implies, PLR carries out OLSR on the individual thresholds to assess how each test location sensitivity is changing over time. In

PROGRESSOR software, the results are presented in the form of a series of histograms for each point tested in the eye. Large bars represent large defects, whilst small ones represent normal vision with colours ranging from white/red (signifying significant deterioration in threshold levels) to grey (indicating no change) to green (signifying significant improvement in threshold levels). A common criterion used to assess whether progression is significant for a single point is a rate less than -1dB/year with a p-value of less than 0.01.

However, the assumption that progression at each test location is linear, and therefore the use of OLSR, has recently been called into question, with pointwise exponential models being proposed instead (Caprioli et al. 2011, Azarbod et al. 2012). Critics of linear regression point to the fact that necessary assumptions of constant variance, normality of residuals (due to floor effects) and independence of measurements are all violated (Pathak et al. 2013). Although, accounting for the latter issue of temporal autocorrelation through mixed modelling has not led to greater model utility (Pathak et al. 2013), two different exponential models have nonetheless been suggested to account for the other two issues: the decay and the non-decay exponential model. The non-decay exponential model is based on the principle that linear measurements of retinal sensitivity (decibels are a logarithmic measure) are proportional to the percentage of ganglion cells lost. This implies that, under the assumption of constant progression, one would theoretically expect rates of loss to appear to increase on the logarithmic decibel scale. The utility of this approach over linear regression appeared to be demonstrated in a study by Pathak et al. (Pathak et al. 2013). Other groups have promoted a contradictory model (the decay exponential model) whereby one would expect rates of decay to decrease over time (Caprioli et al. 2011). One expects that the utility of this measurement is due to floor effects in VF measurements. Tobit linear regression, which is a form of linear modelling that incorporates floor effects, has also been suggested in order to deal with this (Russell & Crabb 2011). Overall, there is not much consistent evidence pointing to one optimal modelling strategy, despite various limitations associated with OLSR. In fact, a recent study comparing various regression methods had OLSR as the best performing in terms of predictive utility in

spite of the fact that assumptions of homoscedasticity (constant variability across sensitivities), normality and independence are violated (Bryan et al. 2013). Overall, linear regression is an undoubtedly imperfect, but overall an adequate compromise of fit and predictive utility in monitoring long-term VF deterioration.

1.5.5 Advantages and Disadvantages of Event and Trend-based analyses

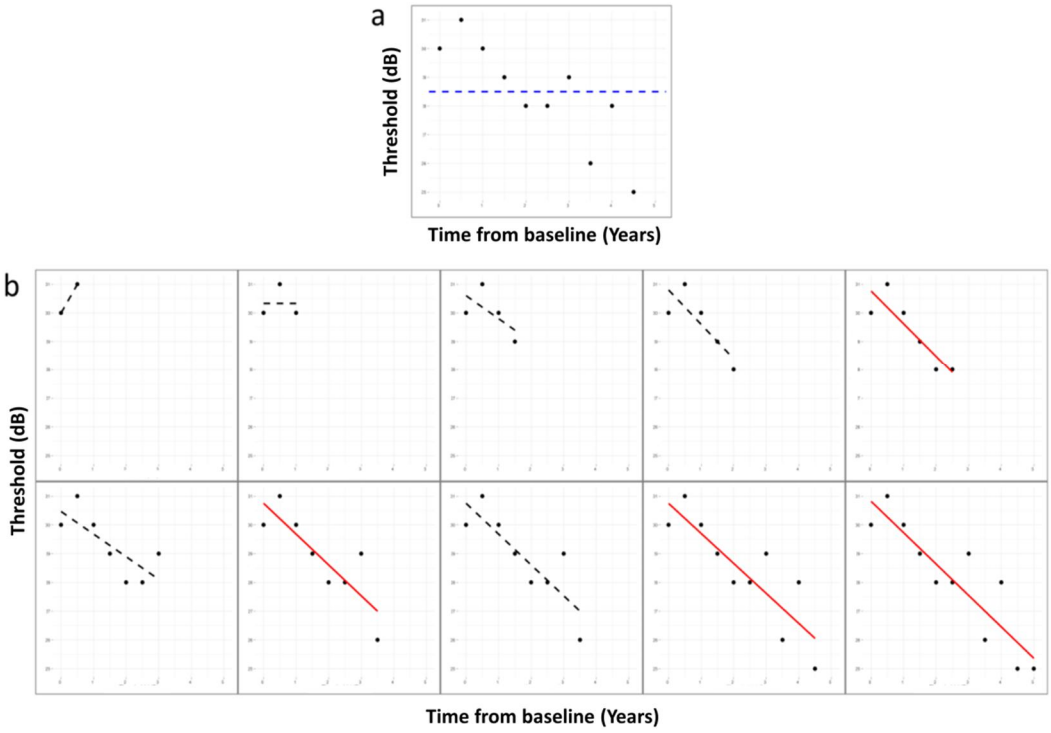

Figure 1.13 - A demonstration of the differences between (a) event and (b) trend based analyses for one point in consecutive visual fields (VF).
In (a), each threshold is compared to the initial baseline derived from averaging two VF measurements (the first two points). If the point is significantly less than the baseline for a stable glaucomatous eye (i.e. below the dotted blue line) for three consecutive VF that point is determined as highly likely to be progressing. Only the baseline and last VFs are used to determine whether progression has occurred.
In (b), for every new VF taken, a regression line is fitted and the significance of it is assessed. If the rate of change is less than 1 dB/year and is significant ($p<0.05$) then that point is deemed to be progressing (solid red line). As can be seen, it is possible for a point to be deemed stable (dotted black line) having been diagnosed as progressing in an earlier field. All VFs are considered in calculating the rate of progression. A version of this figure was published previously in the European Ophthalmic Review (Saunders et al. 2013).

As a result of using all the previous fields in its diagnosis, trend analysis has distinct advantages and disadvantages compared with event based analyses. First of all, event based analyses generally require fewer VFs and less time to produce definitive results, so may detect rapid deterioration in the VF more quickly than this technique. As a result, trend analyses tend to hold a much higher risk of initially falsely diagnosing stable VFs as progressing. For example, in the scenario shown in **Figure 1.13b**, progression would be diagnosed after 3-and-a-half years, but this prognosis would change to not progressing after 4 years (though six months later progression is once again diagnosed). Conversely, trend analysis can also be slower at detecting actual glaucomatous progression than methods such as glaucoma change probability analysis (Nouri-Mahdavi et al. 2007). However, this technique estimates the rate of VF progression, which can be extremely helpful in the context of following patients continuously over a long period of time in clinical practice. Furthermore, with enough fields, trend analyses generally have higher diagnostic sensitivity than event analyses over the same time period (Vesti et al. 2003).

There have been successful attempts to create criteria to help correctly reduce FPs in PLR by imposing limitations on the number of points required to detect progression (Gardiner & Crabb 2002a), yet even these ignore the fact that test locations that are near one another are more likely to be of a similar sensitivity than ones that are further away. Utilising this information could be a big step in improving time to detect progression for analytical methods.

Table 1.1 on page 54 summarises the difference between various methods described here. However, for a more comprehensive review summarising the evidence base on the different methods for assessing glaucomatous VF progression readers are directed to a review by Ernest et al (Ernest et al. 2011). Recently, Nouri-Mahdavi & Caprioli have also reviewed various methods for measuring changes in VF measurements and offer a further exploration of the positives and negatives of approaches to estimating functional, as well as structural, loss (Nouri-Mahdavi & Caprioli 2014). Overall, though flawed in various ways, trend analysis using summary indices or each test location remain the most feasible practical way of monitoring VF progression over time and extrapolating the future VF status of patients.

Table 1.1 – Methods of detecting glaucomatous progression. Acronyms described in Abbreviations Section. This table has been previously published by Saunders et al. (Saunders et al. 2013)

Method	Method of attaining baseline	Method of defining progression	Advantages	Disadvantages	Correction for cataract?	Method Type	Rate of progression calculable?
Linear Regression of MD values	No consensus, but at the very least 3 VFs are required	No consensus, but EyeSuite for the Octopus perimeter defines progression if the previous six MDs are significantly progressing ($p<0.5\%$) (Haag-Streit International 2009)	- Simple	- No spatial consideration of the data - Assumes progression is linear - Long time period required (Smith et al. 1996) - Low sensitivity (Smith et al. 1996) - Affected by cataract	No	Global Summary Trend Analysis	Yes
Linear Regression of VFI values	5 VFs over 3 years required for initial trend in Humphrey Field Analyzer software (Carl Zeiss Meditec 2008)	Assesses significance of slope, and relates rate of progression to how much VFI patients will lose in the next 5 years (Carl Zeiss Meditec 2008)	- Simple - Gives an estimate of the % vision an individual may lose in future	- No spatial consideration of data - Assumes progression is linear - Quite a long time period required - At least 5 VF tests required - Discounts diffuse loss, so may underestimate overall glaucomatous loss	Yes	Global Summary Trend Analysis	Yes
AGIS Method	One VF	A decline in score from baseline equal to 4 "AGIS units" in 3 consecutive tests (AGIS Investigators 1994, Heijl et al. 2008)	- High specificity (Vesti et al. 2003, Heijl et al. 2008, Ernest et al. 2011) - Score testing based on real patient data	- Poor sensitivity (Vesti et al. 2003, Heijl et al. 2008, Ernest et al. 2011) - Cannot determine spatial characteristics of progression - Long time required (Vesti et al. 2003) - Cannot detect progression rate - Can be affected by cataract	No	Pointwise/ scoring event analysis	No
CIGTS Method	Two VFs (Musch et al. 1999, Gillespie et al. 2003)	A decline in score from baseline equal to 3 'CIGTS units' in 3 consecutive tests (Heijl et al. 2008)	- Fast (Vesti et al. 2003) - High specificity (Vesti et al. 2003, Heijl et al. 2008, Ernest et al. 2011) - Score testing based on real patient data	- Low sensitivity (Vesti et al. 2003, Heijl et al. 2008, Ernest et al. 2011) - Cannot determine spatial characteristics of progression - Cannot detect progression rate - Can be affected by cataract	No	Pointwise/ scoring event analysis	No
GCP Analysis	Two VFs	"Likely progression" defined as a reduction in sensitivity (below normal limits) from baseline for ≥3 separate VF points in 3 consecutive tests(Vesti et al. 2003, Heijl et al. 2008)	- Fast (Vesti et al. 2003, Nouri-Mahdavi et al. 2007) - High sensitivity and specificity (Vesti et al. 2003, Heijl et al. 2008, Ernest et al. 2011) - Normal limits defined using real (stable) patient data	- Cannot detect progression rate - Cannot take the diffuse effects of glaucomatous loss into account	Yes	Pointwise event analysis	No
PLR Analysis	No consensus, but at the very least 3 VFs are required	No consensus, but usually a statistically significant ($p<5\%$) decrease of 1dB per year for at least 3 separate VF points	- High sensitivity (Vesti et al. 2003, Ernest et al. 2011) - High specificity (Vesti et al. 2003, Ernest et al. 2011)	- Long time period required (Vesti et al. 2003, Nouri-Mahdavi et al. 2007) - Assumes progression is linear - Can be affected by cataract	No	Pointwise trend analysis	Yes

1.6 Factors affecting time-to-detect progression

The most crucial aim in the monitoring of glaucoma is to diagnose VF progression as quickly as possible. Bearing this in mind, aside from criteria for flagging progression, there are three further factors that affect how quickly quality of life (QoL) threatening deterioration of the VF can be detected: the rate of VF loss, the variability in the VF series (otherwise known as the 'noise') and the frequency at which VF measurements are taken (Chauhan et al. 2008b).

1.6.1 Rate of loss

The rate at which VF is lost is clearly a highly important factor in the ease of detecting progression. The principle behind this is simple: if the magnitude of change in the VF from baseline is higher, then there will simply be greater power to detect a difference whatever method is used to analyse the results. However, obviously in the context of glaucoma, a more rapid rate of loss needs to be detected earlier. It is therefore important to detect deterioration in eyes progressing at a rate, which could very quickly cause visual impairment (Chauhan et al. 2008b).

1.6.2 Noise

Higher variability, or higher levels of noise, in a series makes it more difficult to distinguish actual glaucomatous progression from fluctuations in measurements according to uncontrollable factors. The inevitable result of this is that there is greater uncertainty about progression status lengthening the time required to establish VF loss. Finding ways to minimise or better account for noise is therefore the subject of a lot of research in VFs, although research suggests a 20% reduction in noise is required to make any clinical difference (Turpin & McKendrick 2011).

There are various causes for noise such as technician effect, algorithm, seasonal effects, time of day and patient experience and reliability (Junoy Montolio et al. 2012, Gardiner et al. 2013). However, one major difficulty in measuring VF loss

glaucoma is that levels of noise increase with decreased sensitivity (**Figure 1.14**) (Henson et al. 2000, Artes et al. 2002, Russell et al. 2012a). Progression is therefore easier to detect in test locations with good sensitivity than locations exhibiting glaucomatous damage. Noise reaches a peak at around 10-15dB and decreases towards 0dB due to the fact that sensitivities reach the measurement's lower limit.

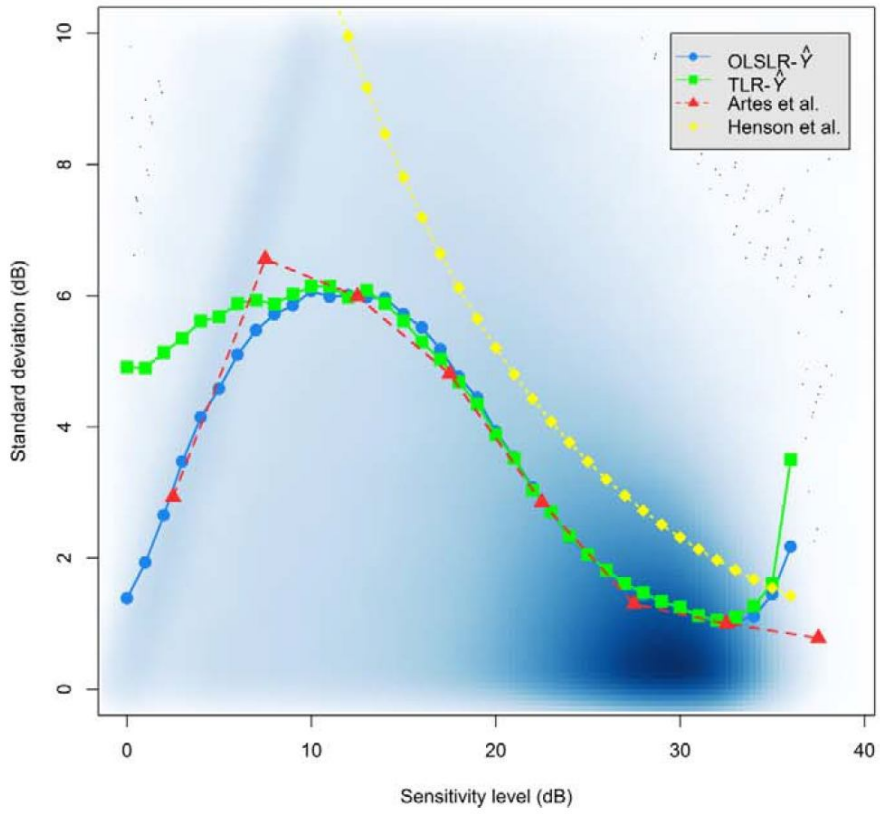

Figure 1.14 – An illustration of how variability changes with sensitivity levels in various clinical studies. The blue background is a density plot of all absolute residuals taken from fitting models to retrospective visual field series in a study by Russell et al 2012a. The blue and green lines represent using least-squares linear regression and tobit regression for these field series respectively. The red and yellow lines reflect the findings of Artes et al 2002 and Henson et al. 2000 respectively. This figure was taken from Russell et al. 2012a.

1.6.3 The Frequency of Visual Field Measurements

A more accurate estimate of the rate of progression is obtained by taking more readings because the underlying signal is more likely to be detected amidst the

variability. Gardiner and Crabb found, using simulated eyes with thresholds decreasing by 2dB per year, that undergoing three VF tests per year was optimal in terms of sensitivity and specificity of detecting progression (Gardiner & Crabb 2002b). However, their estimate of the noise is perhaps too small as it does not take into account the fact that visual thresholds are more variable in areas of lower sensitivity (Heijl et al. 1987, Henson et al. 2000, Artes et al. 2002, Russell et al. 2012a).

One of the most influential papers on this topic written by Chauhan and associates makes suggestions based on the power (the probability of correctly diagnosing a progressing patient) associated with the number of tests taken (Chauhan et al. 2008b). Chauhan et al. suggest, as a minimum, that six VFs should be taken in the patient's first two years of monitoring, before choosing the number of subsequent tests per year on the basis of progression rate and time-scale thereafter (**Figure 1.15**). However, these figures are based on theoretical power calculations (using estimated MD variability of patients in OHTS [Artes & Chauhan 2005]) rather than through the analysis of clinical patient data, so one could argue that this advice is still to be substantiated. In addition, taking more VFs in a period of time will increase the number of FP diagnoses, i.e. decrease specificity (Gardiner & Crabb 2002b).

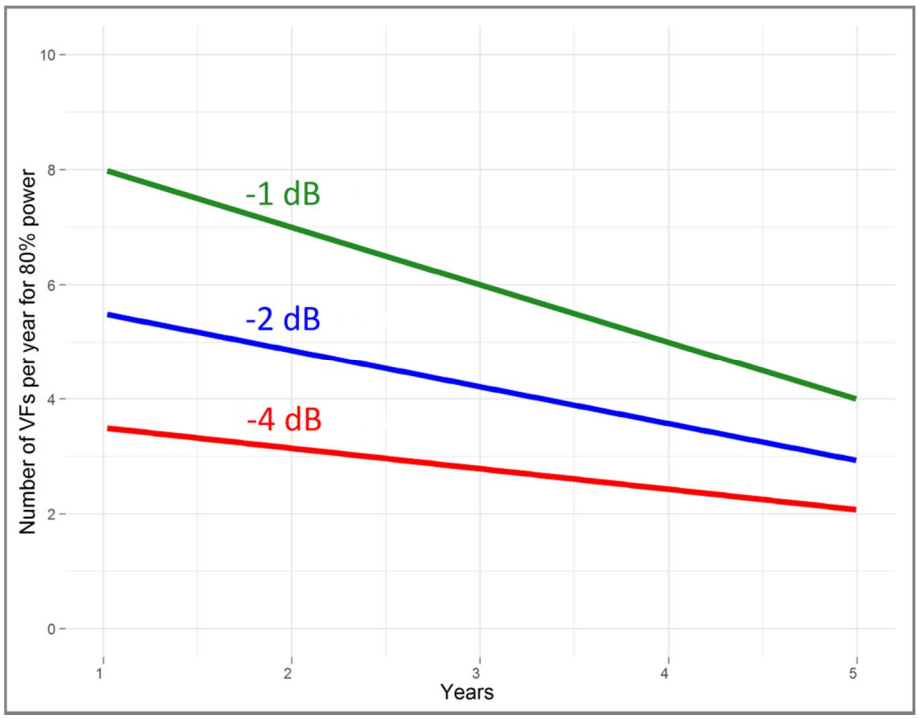

Figure 1.15 – Based on Table 2 in the paper by Chauhan et al. 2008b, this graph shows an estimate of the number of visual fields per year required to have an 80% probability of successfully detecting a progressive change in mean deviation in a given number of years. This figure was first published in the European Ophthalmic Review (Saunders et al. 2013).

Nouri-Mahdavi and his associates conducted a study using clinical trial data confirming Chauhan's basic hypothesis that taking more tests will increase sensitivity of glaucoma progression detection methods and lead to earlier glaucoma diagnosis (Nouri-Mahdavi et al. 2011). They did this by comparing patient progression with all VFs included against that with half of their fields omitted, which roughly translates to comparing bi-annual and annual testing. However, the differences in amount of progression detected between the high and low test frequency groups were less profound than one would theorise based on the findings of Chauhan et al.; the latter publication anticipated that having two instead of one VF test a year would result in a halving of the time required to detect progression (Chauhan et al. 2008b), which is substantially more than differences reported by Nouri-Mahdavi et al. (Nouri-Mahdavi et al. 2011). Part of the conservatism found in Nouri-Mahdavi's study was probably due in part to the fact that many patients had more frequent follow-ups than the 6 month average - the length of time until diagnosis is theoretically halved as the frequency of tests is

doubled, which means that as tests become more frequent the differences in speed of diagnosis become less profound. As a result, taking into account the practical constraints in clinics, Nouri-Mahdavi et al. recommended measuring VFs for patients once every six months.

Perhaps the best recommendation for clinical practice, however, is to vary the intervals between VF tests to optimise detection of progression (Jansonius 2006). One novel approach is to adopt an approach that varies the length of the interval between subsequent tests depending on the outcome of previous test results (Jansonius 2007). It has further been suggested that clustering tests to the beginning and end of a follow-up period (for instance the initial two-year period suggested by Chauhan et al.) rather than spacing tests out evenly can help determine rates of VF deterioration with higher precision (De Moraes et al. 2011b, Crabb & Garway-Heath 2012).

However, in spite of the evidence supporting the importance of performing plenty of tests on patients, in practice, recommended clinical guidelines on frequency of VF testing are largely not followed in clinical practice (Fung et al. 2013, Malik et al. 2013). A key reason for this is likely to be that it is simply not possible given limited resources or capacity (Malik et al. 2013). The inability to achieve this goal, however, represents a serious obstacle to sufficiently monitoring VF progression in clinical practice. Perhaps the key is to stratify patients into those that will benefit from more frequent testing, yet accurate risk profiling of progression awaits further research.

1.7 Visual fields and Visual function

Detecting progression is important, but it is really the impact of VF loss on visual function in everyday life that makes glaucoma an important condition to understand. In fact it has even been suggested that QoL be routinely investigated in clinical practice such is the importance of relating clinical measurements to impact on the patient visual function (Skalicky & Goldberg 2010). Given the fact that the aim of glaucoma care is to ensure that no individual suffers avoidable

visual impairment in their lifetime, it is important to have some idea about what daily activities are affected by glaucoma and the stages of disease at which this happens.

1.7.1 Evaluating patient visual function

Unfortunately, little is known about how VF defects, at different stages of glaucoma, affect patients' abilities to perform everyday visual tasks. This is not only due to the direction of research into this topic being limited (Glen et al. 2011), but also due to the challenges of measuring impairment due to vision specifically in a variety of tasks. Questionnaires (or Patient Reported Outcome Measures [PROMs]) are occasionally used as a measure of QoL, but responses tend to be insensitive and heavily affected by the patient (Bozzani et al. 2012). While prospective performance-based measures, such as the Assessment of Disability Related to Vision (ADREV) (Altangerel et al. 2006), are an objective alternative and have been shown to be more correlated to VF measurements (Richman et al. 2010), it can be difficult to obtain a good quantity of data and there is no guarantee experimental conditions can replicate real-life situations. Meanwhile, analysing retrospective data affords none of the detail of collecting information in an experimental study. A common flaw across methodology is that it is difficult to tell whether performance is just due to glaucoma or affected by other co-morbidities; this is particularly true in the context of the fact that glaucoma is a disease that affects the elderly population. Nonetheless, being able to link the measurements taken in the clinic to what patients visually 'can' and 'cannot do' would be enormously helpful. The best approach is likely to evaluating links between visual function and field loss is to draw evidence from studies using a variety of different methods.

1.7.2 Monocular and Binocular fields

Another large issue when considering patient visual function is that often VF measures are monocular (take measurements from one eye at a time), when it is worth noting when considering VF loss and the effect of this on visual function that patients have two eyes (Asaoka et al. 2011). In the central field in particular there is a binocular summation effect where the VFs of each eye crossover. As a result, damage in the central field of one eye may not impact on the QoL of the patient so long as vision in the other eye is preserved.

A binocular measure of a VF is likely the best way to predict the impact of VF damage on a patient's QoL (Jampel et al. 2002a) although an optimal binocular test that is sensitive to different degrees of VF loss does not exist (Jampel et al. 2002a, Jampel et al. 2002b, Noe et al. 2003). At present, clinicians tend to measure monocular VFs only and binocular VF testing is rarely performed. Therefore, QoL studies often correlate questionnaire responses and task proficiency with the better eye MD (BEMD) or the worse eye MD (WEMD). However, ignoring the vision provided by the other eye can easily lead to overstating the impact of a patient's VF loss on their visual function (Asaoka et al. 2011). This is, in fact, especially important given that there is some evidence that there is a natural tendency for the binocular VF to be preserved in bilateral glaucomatous VF deterioration (i.e. glaucomatous loss in both eyes) (Sponsel et al. 2014).

Ideally, binocular tests should be carried out in order to better assess the impact of disease on an individual's visual function. The only established test for this purpose, however, is the binocular Esterman test (Esterman 1982). This perimetric test procedure randomly presents lights of a fixed intensity of 10dB at 120 separate locations (**Figure 1.16**). The subject taking the test is simply required to indicate whether or not they have seen each point. Unlike SAP, the binocular Esterman test extends into the periphery evaluating the full 160° field in the horizontal (24-2 testing only tests points within 30° of fixation). These locations are not spread evenly; more points are tested in central vision and also in the inferior field.

However, not as many points are tested in the central 30° of fixation as SAP. In addition, although FNs and FPs are tested, FLs cannot be measured using this technique. However, the key disadvantage with this binocular test is that thresholds are not measured. Given the very bright level stimulus presented, the several metrics using this test are not sensitive to change (Harris & Jacobs 1994) and do not tend to correlate well with patient assessment of visual function (Jampel 2001). In fact, the binocular Esterman does not even seem to perform as well as monocular testing at evaluating patient function (Jampel et al. 2002a) Despite these limitations, the Esterman test is nonetheless utilised in the UK in evaluating fitness to drive (Drivers Medical Group 2013).

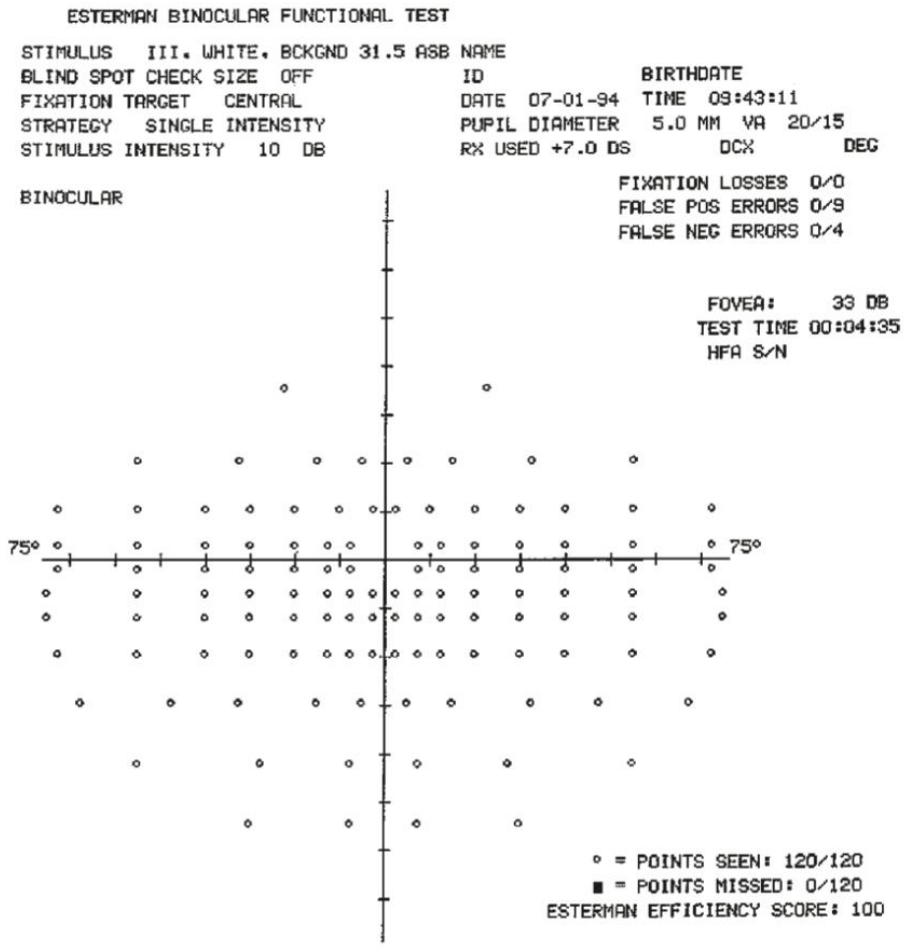

Figure 1.16 – The print-out from an Esterman Test taken from Viswanathan et al. 1999.

Binocular tests can also be taken using 24-2 perimetry and have occasionally performed for some studies (Tabrett & Lathan 2012), but these often are not

utilised in clinical practice as they require a separate test. One simple method of evaluating binocular loss for the central 21° of the VF is to combine VFs from each eye into one binocular VF. The theory behind this is that there are locations of the VF in each eye that correspond the same parts of the binocular VF. There are three main approaches that have been suggested for combining monocular fields. The first such approach, taking the average of corresponding test locations has been shown to have poor correspondence with true binocular testing, so there is little need to mention it further (Nelson-Quigg et al. 2000). The second is known as quadratic summation, which has previously performed well when compared with binocular testing (Nelson-Quigg et al. 2000). It is calculated using the following formula:

$$Binocular\ Sensitivity = \sqrt{Left\ Sensitivity^2 + Right\ Sensitivity^2}$$

However, the most commonly used method combining VF results is the Integrated Visual Field (IVF), also known as the best location method. This approach takes each corresponding location in the VF and compares their sensitivities, taking the highest one at each location (**Figure 1.17**). Although not perfect, a study by Nelson-Quigg and colleagues indicated that this test was closest on average to binocular VF testing, although tended to vary more from Esterman measurements when compared with the quadratic summation method (Nelson-Quigg et al. 2000). Crucially, IVF measurements compare favourably with self-reported patient difficulty (Jampel et al. 2002a, Crabb & Viswanathan 2005, Bozzani et al. 2012) and are also in good agreement with the Esterman in terms of categorising patient legal fitness to drive (Crabb et al. 1998, Crabb et al. 2004, Crabb & Viswanathan 2005, Chisholm et al. 2008). An important limitation of utilising the IVF is that it would not detect patients unfit to drive if they had damage outside of the central 24° that caused them to be legally unfit to drive (see Chapter 2). However, evidence suggests it is relatively uncommon for patients to fail the Esterman test on this criterion alone. In fact, it is more common for patients classified as fit to drive to be classified as legally unfit by the IVF surrogate, even accounting for the fact that the IVF tests more points in the central 20° (Crabb et al. 1998, Crabb et al. 2004, Chisholm et al. 2008). When there is a large discrepancy between the better eye

and the worse eye, the IVF is often very similar to the better eye (Arora et al. 2013), yet the IVF is nonetheless helpful in a number of cases when evaluating patient QoL in bilateral glaucoma.

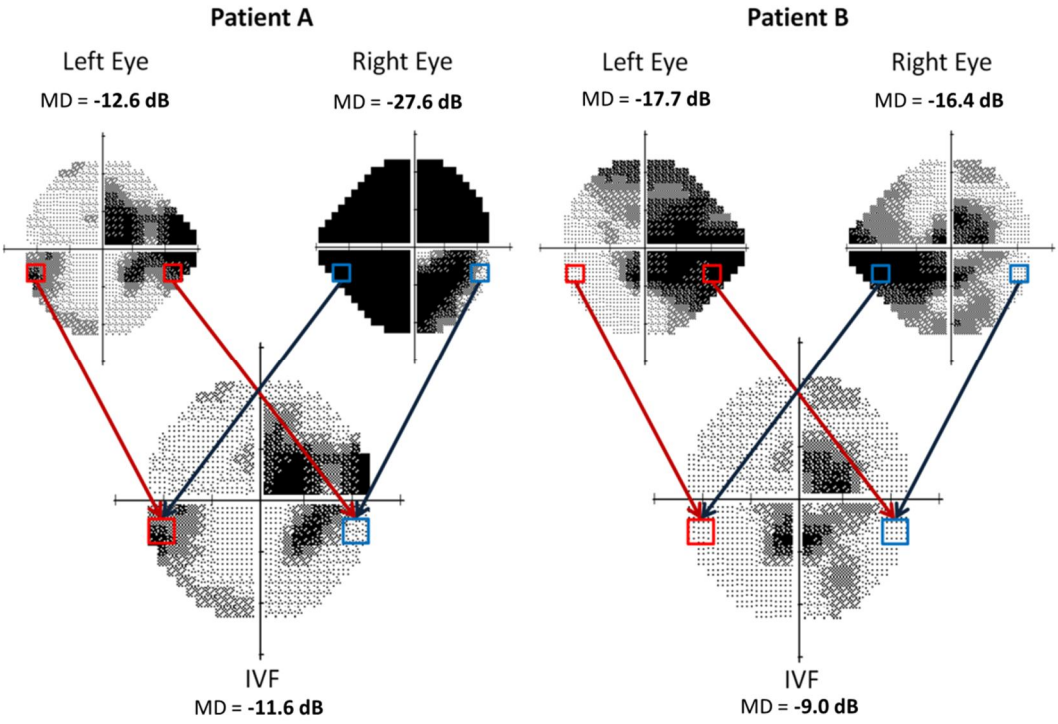

Figure 1.17 - This figure illustrates how the integrated visual field (IVF) is calculated. Corresponding points in the left and right visual fields (VF) are compared and the one with the higher sensitivity is chosen to represent the IVF for that point. The nasal steps are unique to each eye so these are not used in the IVF. The mean deviation (MD) from the better eye can be very similar to the IVF MD in many cases as for Patient A, but can overestimate the severity of binocular damage in cases where damage between the eyes are asymmetric as it is for Patient B.

1.7.3 The effect of glaucomatous loss on visual function

Whatever method has been used to demonstrate VF loss, there is plenty of evidence to suggest that VF deterioration has a profound effect on patient QoL and visual function (Altangerel et al. 2003). Questionnaire responses have not only been able to discriminate well between patients and controls (Goldberg et al.

2009), but have also been able to link visual function and QoL to BEMD (Jampel et al. 2002a), WEMD (Hirneiss et al. 2011, Chan et al. 2013) and binocular measures of VF loss (Parrish, Gedde et al. 1997). This distinction is moreover clearer with greater defect severity (Goldberg et al. 2009, Chan et al. 2013, McKean-Cowdin et al. 2008, van Gestel et al. 2010) although severity is often erroneously based on monocular fields, which do not reflect visual health as well as measures using both eyes. There is evidence that depression score indices are linked to both VF loss and self-reported questionnaire responses (Skalicky & Goldberg 2008).

In addition there is evidence through experimental means that glaucoma affects various tasks that impact on day-to-day QoL. For instance, face recognition (Glen et al. 2012), reaching and grasping speed (Kotecha et al. 2009) and visual search (Smith et al. 2012) have all been experimentally demonstrated to be affected by glaucoma. Richman et al. found stronger links between various experimental tasks as part of the ADREV set of tests, than through PROMs (Richman et al. 2010).

Three activities that have been shown to be important to patients in particular are driving, mobility and reading (Mangione et al. 1998, Aspinall et al. 2008) and there are a number of studies that show these are affected in glaucoma. For instance, not only has VF loss been linked to questionnaire responses on mobility (Noe et al. 2003), but recorded data on the number of falls have also been linked (Black et al. 2011). The issues surrounding mobility for patients are supported by experimental data; Kotecha et al. found that balance could be affected in glaucoma (Kotecha et al. 2012b), whilst speed and success in navigating obstacle courses have also been reportedly impacted in the ADREV tests (Altangerel et al. 2006).

There have also been many studies that have suggested a link between self-reported reading difficulties and glaucoma (Parrish et al. 1997, Gutierrez et al. 1997, Mangione et al. 1998, Lee et al. 1998, Nelson et al. 1999, Janz et al. 2001, Altangerel et al. 2003, Spaeth et al. 2006, Freeman et al. 2008). This has been supported by a number of performance-based studies. The reading of small print task on the ADREV has been shown to be among the most visually demanding tasks in glaucoma patients (Altangerel et al. 2006). Associations between both the speed

of reading out-loud (Ramulu et al. 2009a) and silently (Ramulu et al. 2012) and glaucomatous VF loss have also been shown, even under conditions where differences in acuity should not impact on findings (Ramulu et al. 2009a). These associations do, however, tend to be weak due to the large variability in observed reading speed in patients with glaucoma (Roberts et al. 2005).

Fitness to drive, however, is perhaps one of the most important landmarks to blindness for patients, due to the fact that it can affect an individual's independence and self-sufficiency, something reflected in the fact that studies have linked changes in driving patterns to depression (Fonda et al. 2001). Visual field health is a key component in fitness to drive in the UK (Drivers Medical Group 2013), although the legal criteria behind regulations are based on sparse evidence (Westlake 2000). Although the consensus is that impairment in VF does affect driving (Johnson & Keltner 1983, Owsley et al. 1998a), there are few studies that have found the link between VF loss and risk of accident to be negligible (Burg 1971, McCloskey et al. 1994, Owsley et al. 1998b). However, due to a number of studies being out-of-date and having methodological problems, this is still an area that requires further research. Driving simulations suggest that it is perhaps only when driving tasks become more complex and there is a greater need to concentrate that there is a significant difference in driving ability between glaucoma patients and visually healthy individuals (Vega et al. 2013). Nonetheless, legal fitness to drive notwithstanding, patients with bilateral loss tend to restrict their own driving (Freeman et al. 2006, Ramulu et al. 2009b, van Landingham et al. 2012).

Although it is clear that severity of condition is linked to impact on visual function, at present, it is not obvious for driving or many other tasks at what stage an individual becomes unsafe or unable to participate in a given task. In particular, some association between MD and visual disability would be extremely helpful to give clinicians a general idea of what levels of damage are associated with visual impairment in different day-to-day activities.

1.7.4 Areas of VF and visual function

Assessing patient QoL can be more complicated than simply assessing progression, as the different spatial configurations of defects have different effects when it comes to affecting patient visual function. However, identifying the areas of the VF that correspond with deterioration in QoL can be problematic, generally due to the fact that, in addition to it being very difficult to assess patient performance, it is a substantial challenge to get a good sample of patients with specific damage configurations. Nonetheless, there have been various studies that have attempted to look in more detail at how various defect types affect visual function and it is clear that different parts of the VF are important for different daily activities.

Certainly defects close to fixation are debilitating, being consistently linked to more negative questionnaire responses (McKean-Cowdin et al. 2008, van Gestel et al. 2010). This is especially true in reading, which has been shown to be particularly dependent on the central field (Whittaker & Lovie-Kitchin 1993, Tabrett & Lathan 2012). However, research observing visually healthy people with simulated VF loss suggests that individuals have more difficulty adapting to inferior field loss in terms of slowed reading speed than superior, nasal or temporal defects (Cummings, Rubin 1993).

A recent UK study aiming to explore what parts of the binocular VF may correspond most closely with day-to-day difficulties for patients with visual impairment, including a small number of glaucomatous patients, suggested that whilst the central 5 degrees of the VF was particularly important in reading, the periphery of the central field (10-30° from fixation) was most associated with mobility (Tabrett & Lathan 2012). A study from Black et al based in Queensland further suggest that it is inferior defects rather than superior ones that are most closely linked to risk of falling (Black et al. 2011). The inferior paracentral part of the VF has additionally been linked with reaching and grasping tasks (Kotecha et al. 2009), whilst face recognition, like reading, predominantly seems to be dependent on the health of the centralmost VF (Glen et al. 2012).

An activity such as driving is more complex, with any loss potentially related to some facet of the activity. There is some evidence that central field loss does seem to be more critical than loss in the peripheral field (Kooijman et al. 2004), while, in addition, the superior hemifield seems to be more important than the inferior (Vega et al. 2013, Glen et al. 2014). However, it is unclear whether or not there is a difference between defects being in the left or right side of the field. A Canadian study (driving on the right) noted that diffuse loss in the right hemifield, as well as focal loss in the left hemifield were both associated with inadequate performance in on-road assessments (Racette & Casson 2005), whilst a London-based study (driving on the left) simulating damaged VFs onto the UK hazard perception test did not note any clear difference in performance regarding the left or right location of the VF defect (Glen et al. 2014). However, due to large intra-individual variability in terms of how defects impact on driving, being able to assess driving ability based on configuration of the defect is still a long way away and it would perhaps still be best to base driving ability on assessed practical performance (Racette & Casson 2005).

One of the reasons for the difficulty of relating VF health to driving proficiency is not only the complexity of eye movements involved in driving and the difficulty involved in measuring fitness to drive, but also the fact that patients can adapt to their defects. For instance, Crabb et al. note that patients taking the hazard perception test attempt to compensate for VF losses by making more saccades (Crabb et al. 2010), although this was not observed in another study that used a driving simulator (Vega et al. 2013). A study by Coeckelbergh et al. also attempted to look at the relationship between eye movements and driving in the hope that patients could adapt to their VF loss and become safer drivers with training. Whilst they found correspondence between search tasks and numbers of fixations, no eye movement parameters were found to correspond closely with a practical driving exam (Coeckelbergh et al. 2002). It was further shown that many of this cohort could not be made fit to drive even after training (Kooijman, Brouwer et al. 2004).

One of the most interesting studies on the subject, from Murata et al., attempted to use the machine learning orientated Random Forest method to look at which

areas of the VF are linked to various daily activities, linking VF locations to questionnaire responses (Murata et al. 2013). Although, the worse eye acuity was still the most important predictor of functional loss, the study found the peripheral superior left and inferior left hemifields were important for reading, with the peripheral, mid-peripheral and para-central inferior regions important in walking, whilst the superior peripheral was most important in the set of questions related to taking public transport. Broad scattered areas of the VF are related to dining (**Figure 1.18**). This study is perhaps the most interesting illustration of how VF defects of various types can impact on the lives of patients suffering them in such differing ways.

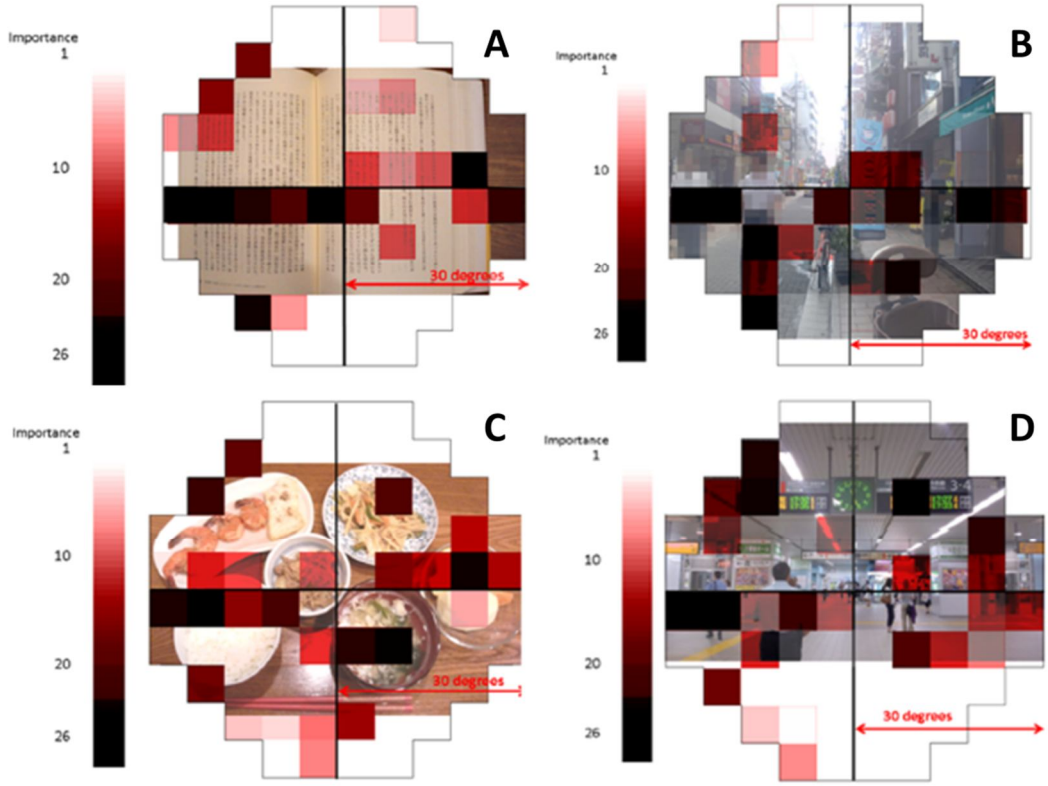

Figure 1.18 – Findings taken from the Murata et al. 2013 study showing the different parts of the visual field (VF) important in A) Reading, B) Walking, C) Dining and D) Going out. Darker colours correspond with more important locations in the VF. Image was taken from Murata et al. 2013.

1.7.5 Visual fields and life expectancy

Evaluating the level of VF loss is important in determining the current status of an individual and determining progression rates are highly important in calculating risk of future blindness or visual impairment. However, progression rates are contextless without taking into account how long the patient is likely to live. It is thus important to use information on VF progression to assess whether a patient requires further intervention or not dependent on whether the condition is progressing quickly enough to have a tangible effect in their expected lifetime, thus determining whether treatment is saving sight years (Heijl 2013). One intuitive means of doing this is to use patient life expectancies as Wesselink and colleagues recommend (Wesselink et al. 2011). However, although life expectancies can give us a good idea, the length of time patients will live for is always uncertain and what severity of damage corresponds to different daily activities is yet unknown (**Figure 1.19**) (European Glaucoma Society 2008).

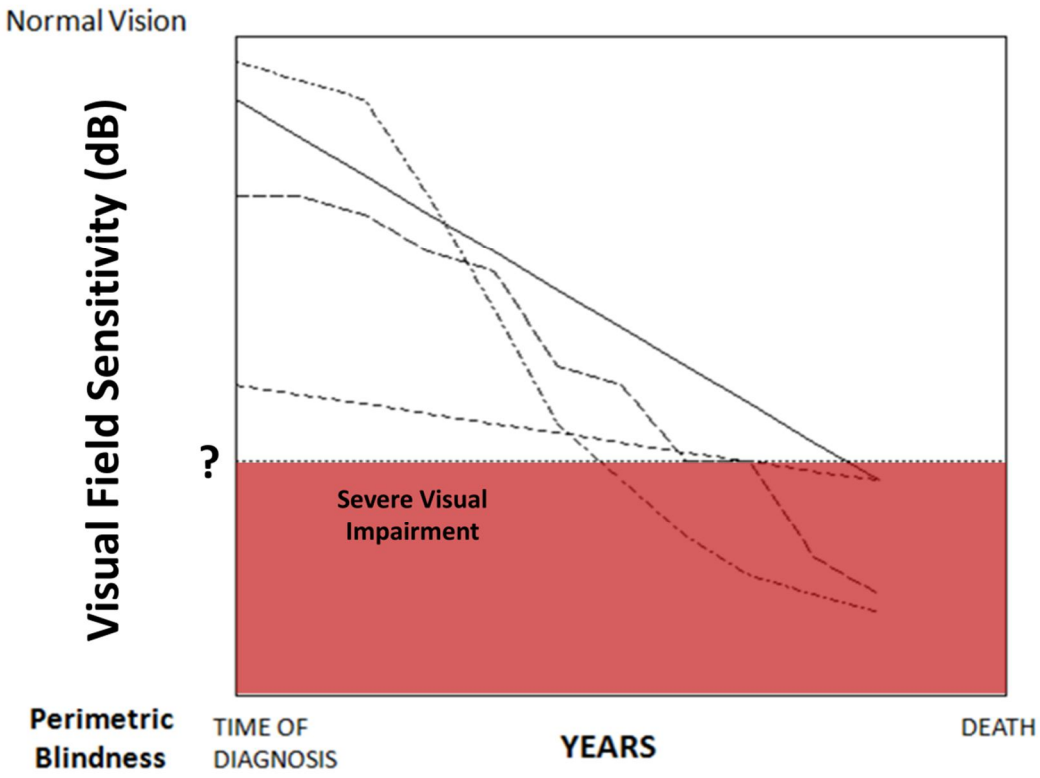

Figure 1.19 – An illustration of the conundrum associated with monitoring visual field loss over time. The aim of glaucoma management is to prevent patients from reaching a state of severe visual impairment within their lifetime, yet it is not clear when patients have progressed to blindness. This figure is based on an image from the European Glaucoma Society Guidelines (European Glaucoma Society 2008).

1.8 Objectives

It is clear that there are still plenty of unknowns when it comes to how the measured VF can impact on QoL and furthermore how best to monitor it over time to prevent patients from reaching a stage of visual impairment. This thesis aims to explore a few of these questions.

Chapter Two begins by exploring the impact of VF loss on an activity that has a large impact on individual QoL, meeting the legal requirements of being fit to drive. The aim of this chapter is to investigate what levels of central VF loss (24-2 pattern)

corresponds to failure to the VF component of legal fitness to drive in the UK measured using a surrogate of the Esterman VF test.

Chapter Three then explores how many individuals with glaucoma are progressing at a rate that puts them at risk of progressing to this level of visual impairment. To do this, median life expectancies based on patient age are taken from the Office of National Statistics (ONS) and rates of VF loss are projected to predict future visual function. In this modelling exercise, numbers of patients anticipated to progress to visual impairment and statutory blindness are analysed.

Having established the number of patients at risk of visual impairment and the importance of monitoring patient VF loss effectively to find those likely to progress quickly, Chapter Four then looks at whether there is a difference between using the two testing algorithms SITA Standard and SITA Fast in the context of precision and the impact on time needed to detect VF progression.

Finally, Chapter 5 investigates whether risk factors at baseline can be utilised to detect those most at risk of reaching a state of visual impairment. The utility of one current proposed published risk calculator for predicting future progression rates in particular is described and estimated.

The final chapter summarises the work in this thesis and looks briefly at other research currently in the process of being undertaken along with future research projects.

1.9 Data

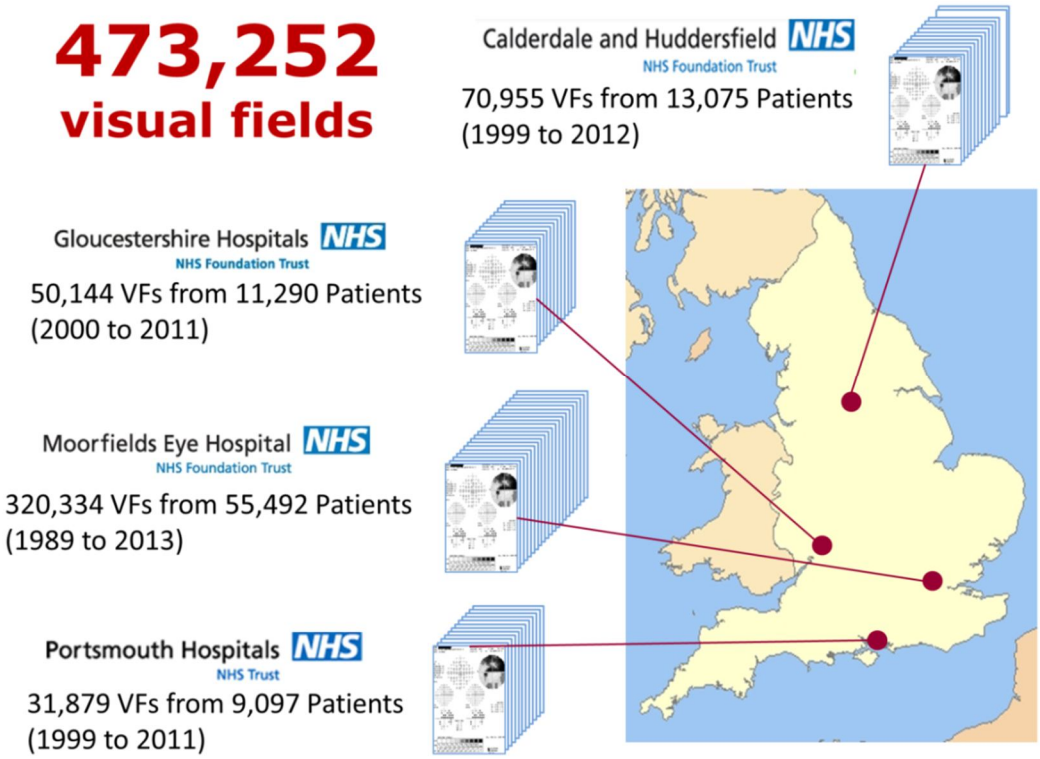

Figure 1.20 - The locations of the four clinical centres in England where visual fields used in this thesis were collected

In order to explore the topics outlined in this thesis above, a wealth of data has been utilised. This research uses real-life clinical data from 473,252 anonymised VFs taken from four NHS glaucoma clinics across the UK. These centres include Moorfields Eye Hospital in London (320,334 VFs), Cheltenham General Hospital Gloucestershire Eye Unit (50,144 VFs), Queen Alexandra Hospital in Portsmouth (31,879 VFs) and the Calderdale and Huddersfield NHS Foundation Trust (70,955 VFs) (**Figure 1.20**). The data was collected with patient consent and recorded onto a Medisoft database. Access to the data was granted by the Caldicott guardians at each centre. The study adhered to the Declaration of Helsinki, was approved by a research ethics committee of City University London and all anonymised data were transferred to a secure database. The inclusion criteria varied from study to study and will be looked at in more depth in the coming chapters.

Chapter Two: Visual field measurements and legal fitness to drive – deriving practical landmarks for visual field disability in glaucoma

Visual field measurements are the only current means of directly assessing what a patient with glaucoma can see and are therefore essential in monitoring functional disease progression to aid clinical decision making in preventing blindness. Yet even in those patients where risk of blindness is averted, glaucoma can still cause 'visual disability' that impacts on QoL (Ramulu 2009). For example, an association has been demonstrated between VF loss and self-reported QoL among patients who were *unaware* of their glaucoma at the time of QoL interview (McKean-Cowdin et al. 2008). There are many ways of summarising VF damage, but how these measurements relate to visual impairment remains largely unknown and requires further research (Glen et al. 2011). Though many attempts have been made to categorise VF defect severity and grading criteria (Susanna Jr. & Vessani 2009), to date these criteria have not been set with stages of visual impairment in mind. For instance, none of these staging criteria can be linked to legal fitness to drive (LFTD) in the UK. This is in spite of the fact that driving cessation has been shown to be one of the most incapacitating consequences of glaucoma (Ramulu et al. 2009b, van Landingham et al. 2012) and has been shown to be a major negative QoL "landmark" for patients (Bhargava et al. 2006) being linked to depression (Fonda et al. 2001).

In the UK, the binocular Esterman VF test (Esterman 1982) is used to assess LFTD by the Driving Vehicle and Licensing Agency (DVLA) (Drivers Medical Group 2014). The 120 point Esterman tests a fixed stimulus threshold of 10dB at 120 test locations. To satisfy the VF component of LFTD, current regulations stipulate that the person should have "no significant defect in the binocular field which encroaches within 20˚ of fixation", and the ability to use "a field of at least 120˚ on the horizontal" meridian (Drivers Medical Group 2014). The definition of a significant defect is debateable, but the current definition is: "a cluster of four or

more adjoining [failed (missed)] points that is either wholly or partly within the central 20 degree area" or "loss consisting of both a single cluster of three adjoining missed points up to and including 20 degrees from fixation, and any additional separate missed point(s) within the central 20 degree area" (Drivers Medical Group 2014). In other words, four contiguous missed points or three plus an additional missed point within 20° of fixation. The peripheral vision regulations are more vague, stating that patients can pass with "up to three adjoining missed points...lying on the horizontal meridian" or a "vertical defect of single point width but of any length, [that] touches or cuts through the horizontal meridian" (Drivers Medical Group 2014). A clinician may only advise a sight impaired patient to inform the DVLA of their status with the onus being on the patient to do so. However, clinicians are allowed to breach confidentiality when their advice is ignored and the patient is considered a danger to himself and others in a vehicle (Royal College of Ophthalmologists 2013). In practice, clinicians have been shown to be poor at advising glaucomatous patients with binocular defects to contact the DVLA (Puvanachandra et al. 2008).

The aim of this chapter is to provide a first step towards relating VF summary indices from the 24-2 test pattern and LFTD. Specifically, this section analyses the relationship between the MDs of patients' monocular VFs and an IVF surrogate measure for the VF component of legal fitness to drive in order to evaluate what levels of VF damage are associated with losing one's license in the UK. The work in this chapter has formed a paper published by the *British Journal of Ophthalmology* (Saunders et al. 2012a). Richard Russell (RR) and David Crabb (DC) were joint authors of the paper; RR contributed to the data analysis and the work was directed by RR and DC. I prepared the data, performed all the analysis, wrote the paper and produced all results and figures; RR and DC edited and revised the paper. This work was also presented as a read paper at the United Kingdom and Eire Glaucoma Society Meeting in Edinburgh, UK on 6-7th December, 2012. All data utilised for this chapter was from Moorfields Eye Hospital, as this was the only data available at the time of the study.

2.1 Methods

This study retrospectively investigated 68,099 anonymised VFs collected from 8,252 patients visiting the Glaucoma service at Moorfields Eye Hospital between 1997 and 2009. The study adhered to the tenets of the Declaration of Helsinki and was approved by research ethics/governance committees of the participating institutions. All data were anonymised and transferred to a secure computer database at City University London. In all cases VF testing was carried out with the HFA using the 24-2 test pattern with a Goldmann size III target and the SITA Standard testing algorithm. VF tests with FP or FN rates above 30% or FLs greater than 20% were discarded. Patients were only included if both eyes had VFs with an MD flagged as worse than -2.07dB ($p<5\%$) in order to have a sample representative of patients who might be referred for an Esterman test. Finally, only a patient's most recent monocular right and left eye VF tests (performed on the same day) were included. In total, 5,208 VFs from 2,604 patients were examined.

2.1.1. Estimating legal fitness to drive using the IVF

The IVF was utilised as a surrogate measure to assess whether a patient would be at risk of passing or failing the Esterman test criteria for LFTD. This method has been shown to give very good agreement with the Esterman concerning the UK VF criteria for LFTD (Crabb et al. 2004, Chisholm et al. 2008). The derived Binocular VF thresholds were dichotomised into groups with thresholds ≥10 dB and thresholds <10dB representing whether a patient would see or miss a point in the Esterman test, respectively. The number and location of these points were assessed and compared with current DVLA standards to categorise patients as those that would satisfy the VF component for LFTD (Crabb et al. 2004). Since the number of points tested in the inner 20° of the IVF is 32 instead of 24 in the Esterman, 6 or 4+1 (four aligned defective points within or partly contained within 20˚ of fixation plus at least one additional defective point also within 20˚ of fixation) clustered points were required for a patient to be deemed to legally unfit to drive, as has been established by Crabb et al. (Crabb et al. 2004). As the IVF using 24-2 measurements only extend to a maximum of 21 degrees from fixation, it was not

possible to assess LFTD according to regulations applied to the peripheral VF. This process was part-automated, using code purpose written in the open-source statistical environment R (R Development Core Team 2014) to filter out patients having too few defects to fail, and then checked manually for cases where the definition required adjoining points. MDs were calculated using the PROGRESSOR software and the BEMD (i.e. the eye with the better MD) and the WEMD were determined for each patient. It was assumed that the prevalence of glaucoma patients who would fail the surrogate test in this study was similar to the proportion of the total population of patients with existing bilateral functional loss who would do likewise.

2.1.2. Analysis

BEMD and WEMD were compared using Receiver Operating Characteristic (ROC) curves, which represent the specificity and sensitivity (Altman & Bland 1994a) of the variables to correctly classify legally fit to drive patients (LFTDP) and legally unfit to drive patients (LUTDP) at different thresholds, respectively (Altman & Bland 1994c). The area under the curve (AUC) represents a criterion-free measure of the curves' comparative diagnostic power. Confidence intervals (CI) for the AUC were calculated using DeLong's method, which is an asymptotically exact method of estimating CIs in this context (DeLong et al. 1988).

It was of interest to see whether developing a model using data from both eyes would give an improved result, so a logistic regression model was used to combine information from both eyes and fitted values were produced from which an ROC curve could be produced to test the hypothesis that utilising information from both eyes would improve diagnostic ability.

"Probability of failure" (PoF) was defined as the proportion of LUTDP below a given MD threshold (this is also known as the positive predictive value). The PoF, as the main outcome measure of the study, was estimated for all MDs alongside the sensitivity and specificity to assess diagnostic coverage, as well as the negative predictive value (NPV), which is the probability of a patient being legally fit to drive

given an MD is above the threshold (Altman & Bland 1994b). Statistical analyses were carried out in R and ROC curves were generated using the software package pROC (Robin et al. 2011).

2.2 Results

Some demographic information for the 2,604 patients investigated is shown in **Table 2.1**. Three-hundred and seventy patients were predicted to fail the Esterman test based on their IVF (failure prevalence equal to 14.2%).

Table 2.1 - Patient information and Mean Deviations (MDs) of the cohort LFTDP indicates those patients who are 'legally fit to drive' according to the surrogate IVF test while LUTDP indicates those patients who are 'legally unfit to drive' according to this test.

		Median	Interquartile Range
Month of test		December 04	January 02 to October 07
Age (years)		69	58 to 76
Better eye MD (dB)	**Overall**	-4.7	-3.1 to -7.8
	LFTDP	-4.2	-3.0 to -6.2
	LUTDP	-14.0	-10.8 to -16.9
Worse eye MD (dB)	**Overall**	-9.4	-5.5 to -15.3
	LFTDP	-8.1	-5.1 to -12.7
	LUTDP	-19.4	-15.6 to -22.9

2.2.1 ROC curve Analysis

In **Figure 2.1** an ROC curve is fitted to the data for the BEMD; the area under curve

(AUC) statistic is equal to 96.2% (95% DeLong CI: 95.4 to 97.1%). The ROC curve for the WEMD is not shown, but, as expected, this metric had poorer diagnostic utility (AUC: 89.2% [95% CI: 87.8-90.6%]). This is reflected in the fact that the ROC curve for the WEMD in **Figure 2.1** falls outside the 95% CI for the BEMD across specificities, suggesting that the BEMD is the better diagnostic of legal fitness to drive.

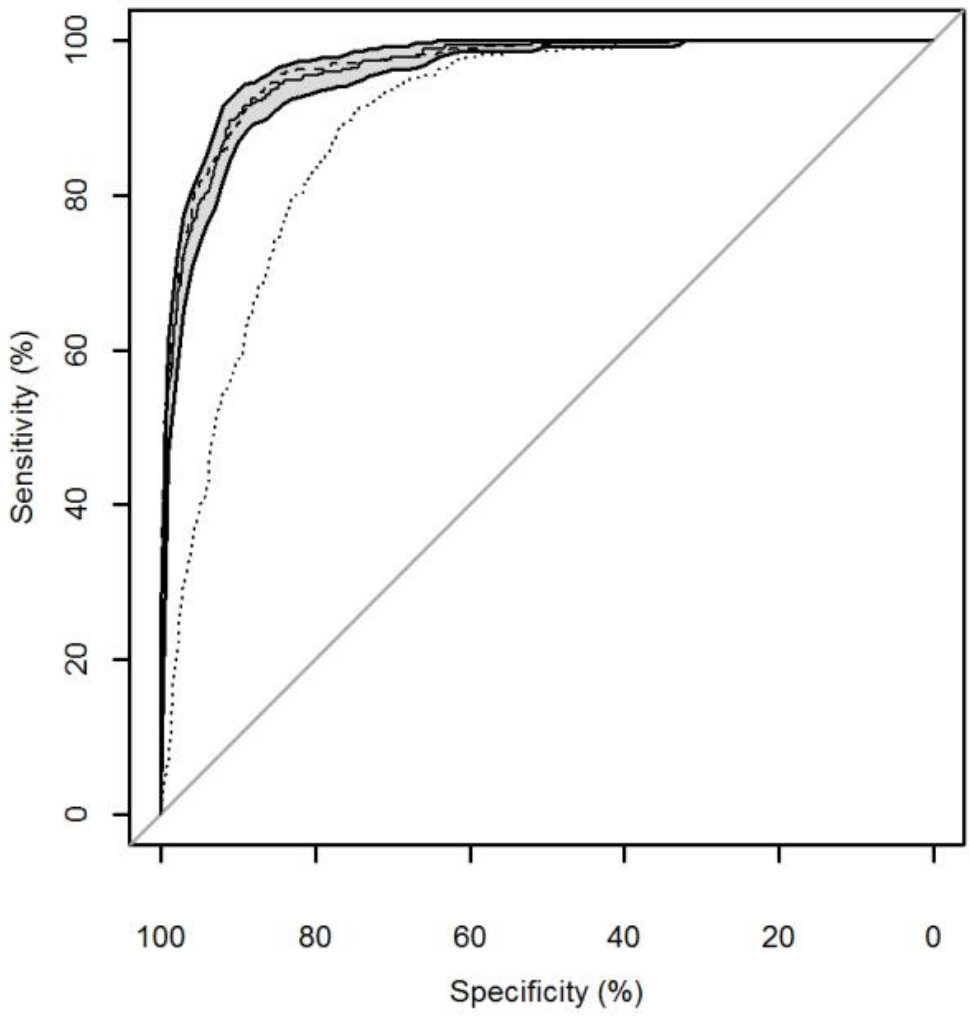

Figure 2.1 - A receiver operating characteristic (ROC) plot for using different summary measures for predicting the IVF surrogate measure of legal fitness to drive. The grey diagonal line represents the random-guess line (line of no discrimination), the solid line the ROC curve for the better eye mean deviation (MD) with the grey band is its 95% CI generated using bootstrapping. The lower dotted line is the ROC curve for the worse eye MD, while the higher dashed black line is the ROC curve derived from the logistic model of worse and better eye MDs.

The fitted logistic regression model and parameter estimation for the BEMD and WEMD can be summarised as:

$$\eta = \ln\left(\frac{probability\ of\ failing\ IVF\ test}{1 - probability\ of\ failing\ IVF\ test}\right)$$

$$\eta = -7.86 - 0.10(WEMD) - 0.53(BEMD)$$

This equation strongly suggests that the better eye MD drives the model given that its effect size is over five times that of the worse eye MD (95% CIs: -8.60 to -7.18, -0.14 to -0.06 and -0.59 to -0.47 for the intercept, WEMD and BEMD coefficients respectively). Fitted probabilities were generated for each η value and an ROC curve for the model's fitted probabilities was produced (**Figure 2.1**); it is clear from visual inspection that there is no significant benefit from using the logistic model over the BEMD (AUC: 96.4% [95% CI: 95.6-97.2%]).

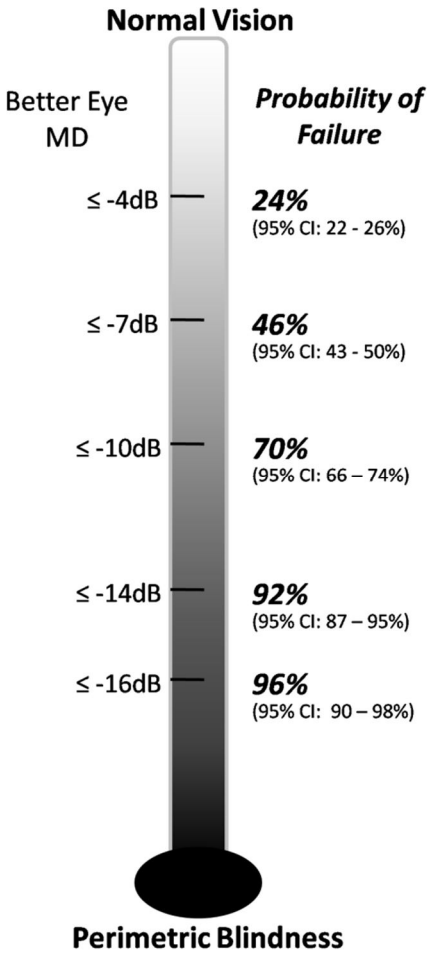

Figure 2.2 - A schematic showing the relationship between defect levels (better eye mean deviation) and the probability of failure of the surrogate Esterman test with 95% confidence intervals (CI).

2.2.2 Calculating the Probability of Failure

PoF is given for some BEMD values in **Figure 2.2**. A BEMD threshold of ≤-7.0dB serves as a promising referral statistic with sensitivities and specificities of 95% and 82%, respectively; these imply a good diagnostic coverage rate but the PoF is quite low (46%). In other words, 46% of patients in this sample with a BEMD ≤-7dB would be legally unfit to drive according to the surrogate Esterman test, which constitutes 95% of everyone in this study who would fail the test. Nevertheless, less than 1% of patients with a BEMD ≥-7dB

would fail the surrogate Esterman (NPV: 99%). A BEMD threshold of ≤-10dB provided less diagnostic coverage than the -7dB threshold (sensitivity and specificity equal to 80% and 94%, respectively), but PoF was higher (70%), thereby offering a good compromise between coverage and accurate diagnosis. In this case, 70% of the patients in this sample with a BEMD of -10dB or worse would fail the surrogate Esterman test, which constitutes 80% of all who would fail. PoF was very high (over 90%) when a patient's BEMD was ≤-14dB.

2.3 Discussion

The aim of this study was to establish whether there is a meaningful relationship between VF MD and the IVF failure to drive threshold. The statistics given in **Table 2.1** suggest that BEMD was a good predictor of whether a patient is legally fit to drive and this is supported by Figure 1, which illustrates the diagnostic performance of the BEMD to classify whether a patient will pass or fail the surrogate Esterman test. A BEMD value of -7dB appears to be a clinically useful landmark, and could serve as a reminder to refer a patient for an Esterman test if this has not been advised already. A BEMD of -14dB or worse could be a useful threshold as a clinical trial endpoint for "visually disabling" VF damage as it suggests that a patient is very likely to fail the DVLA criteria for the Esterman test. Of course, no threshold is a perfect diagnostic as **Figure 2.3** demonstrates but, interestingly, many of the 'false-positive' patients in the study (including the one shown in **Figure 2.3**) would have failed the surrogate Esterman test if just one more failed point was present in their IVF, raising important questions about the current VF criteria for LFTD, discussed elsewhere (Westlake 2000).

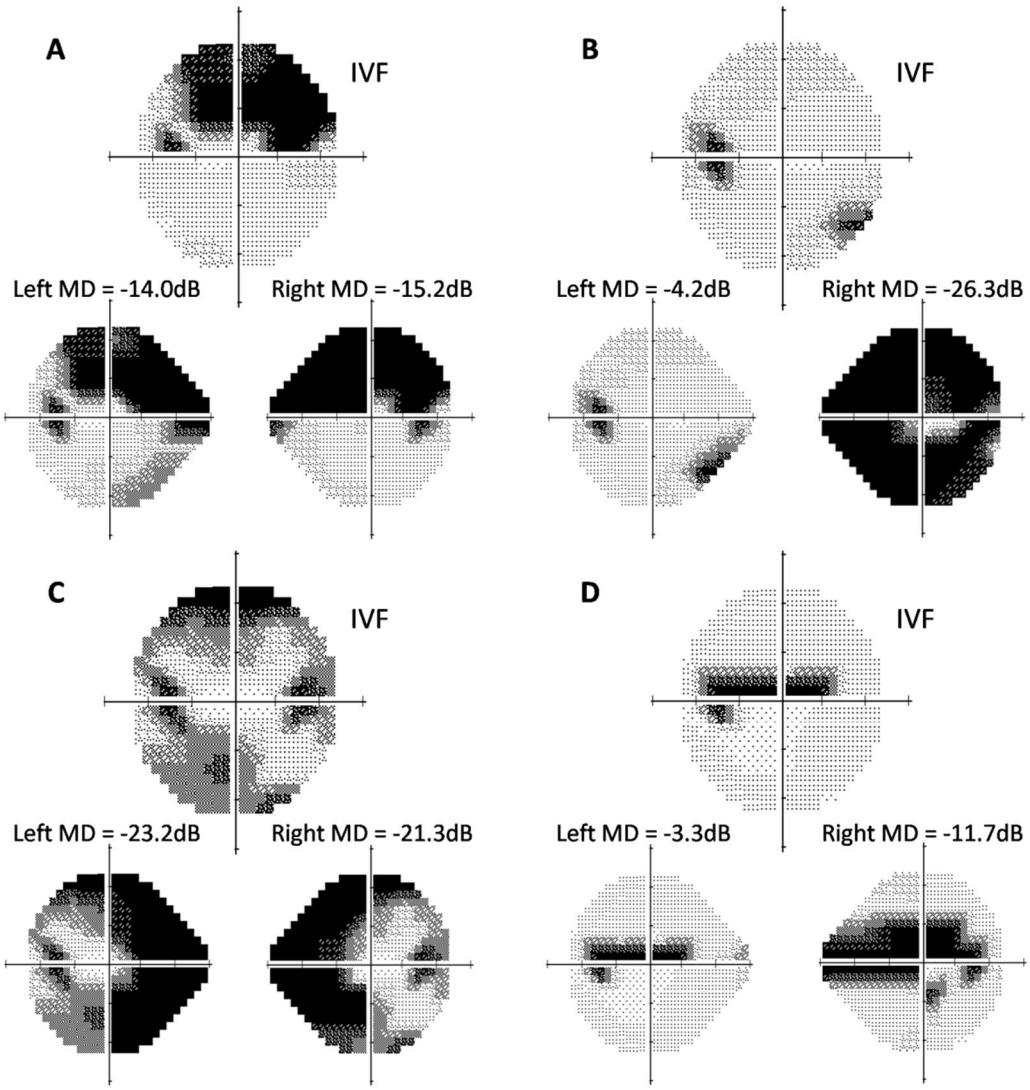

Figure 2.3 - Examples of 'true-positive' (A), 'true-negative' (B), 'false-positive' (C) and 'false-negative' (D) patients. Patient (C) has advanced visual field (VF) damage in both eyes; however, the damage is spatially asymmetric, and so the patient passes the surrogate Esterman test. Conversely, Patient D has less severe VF damage in both eyes, but the damage is spatially symmetric, and so the patient fails the surrogate Esterman test. Integrated VF and greyscales were generated in R.

Unsurprisingly, this study has limitations; most notably, the IVF is not the same as the Esterman test. In particular, the possibility exists that patients with healthy central fields may have binocular peripheral loss that contravenes legal fitness to drive requirements, as no points beyond 21° from fixation are utilised in the IVF.

However, previous research suggests that the 'better sensitivity' method for generating the IVF represents a good measure of the 'true' binocular VF in patients with glaucoma (Jampel et al. 2002a, Nelson-Quigg et al. 2000). Furthermore, the IVF has already been demonstrated as practical for assessing LFTD compared to using monocular VFs (Owen et al. 2008) and a strong correspondence between IVF and Esterman test results has been shown in glaucoma (Crabb et al. 2004, Chisholm et al. 2008, Crabb et al. 1998) with a minimum level of agreement equal to 88% in all cases. Disagreements between the two tests can be attributed to the higher sensitivity of the IVF due to its higher density of test points in the central and superior regions (Crabb et al. 2004), including the four points closest to fixation, which the Esterman does not test. This suggests that the sensitivity and NPV of the better eye MD to assess Esterman measured LFTD is even higher since there is evidence the IVF surrogate test may be more stringent than the actual Esterman test (Crabb et al. 2004). Furthermore, the IVF offers advantages over the Esterman test because it is possible to measure a patient's fixation (Kotecha et al. 2008) and it is based on a threshold test rather than a supra-threshold test.

Another limitation concerns the sample being representative of patients with binocular VF damage that would be considered for referral to assess LFTD. This is pertinent given that it is possible to attain almost perfect sensitivity and specificity by supplementing the sample with people without significant bilateral VF loss. This bias has been minimised by only including patients with VF defects in both eyes, as measured by MD, thereby representing the population of patients that should be referred according to current guidelines (Royal College of Ophthalmologists 2013). More than 2,500 patients with a range of VF defect severities were analysed; a sizeable sample to allow important conclusions to be drawn. However, it is important to note that the results are not applicable to patients with VF defects from neurological conditions (Chisholm et al. 2008) and can only be related to glaucomatous patients with existing VF defects (i.e. not glaucoma suspects or ocular hypertensives).

The IVF test results are not directly equivalent to the Esterman test, which is the means for testing LFTD, and this paper does not promote the usage of summary

statistics over the Esterman test. In particular, MD discards important spatial information about a VF defect, which is an important feature of glaucoma when assessing LFTD, but the index *is* readily available to clinicians from their routine clinical follow-up of patients. Perhaps, it is worth echoing suggestions that perimetry analysis software could include analyses such as the IVF (Owen et al. 2008). For now, results from this study can potentially act as a marker to aid clinicians in making treatment decisions that will prevent the QoL of a patient from being affected in their lifetime, as well as serving to remind the clinician to remind patients with binocular VF damage to notify the DVLA. This report could be a step towards bridging the gap between a patient's real life disabilities and their VF damage, and such landmarks may prove particularly useful as clinical trial endpoints or for audit purposes.

The results of this section, thereby, provide clinicians, who only have access to monocular SAP measurements, evidence-based justification for assessing a patient's visual disability, aiding clinical decisions. They do not promote the use of summary measures as a replacement for standards in measuring the VF component of LFTD, but instead aim to encourage a framework for interpreting VFs in the context of a patient's visual disability rather than the current precedent which is devoid of a reference standard.

Having some impression of what severities of VF loss correspond to loss in QoL is useful in allowing clinicians to prevent patients with glaucoma from reaching impairment. However, it is also important to take into account that glaucoma is a long-term progressive condition and patients can live for years with the condition, with VFs deteriorating over time. It is imperative to ensure that disease progression is controlled so that patients do not reach a stage of visual impairment in their lives, whilst not burdening them with overtreatment. In the next chapter, I investigate how many patients in clinical practice are progressing at a rate of loss likely to lead to these levels of damage within their predicted remaining lifetime.

Chapter Three: Examining visual field loss in patients with glaucoma during their predicted remaining lifetime

The previous chapter established the probability of being legally unfit to drive (a major determinant of patient vision-related QoL) across a series of MD measurements. In particular, it was established that VF loss worse than -14dB in the better eye was associated with an estimated 9 in 10 chance being legally unfit to drive in the UK. Another key landmark for patients is becoming registered as legally blind or severely sight-impaired (SSI) in the UK. The definition for becoming legally SSI is based upon both the visual acuity and VF; patients can be diagnosed as SSI with an corrected visual acuity of worse than 3/60 without field loss or with a much reduced field of vision with healthy acuity. In the US, statutory (legal) blindness is typically defined as best corrected visual acuity of 20/200 or worse, but also includes a definition based on VF limitation (US Social Security Administration 2011). In particular, the widest diameter of the central VF, in the better eye, must subtend an angle at least 20 degrees (i.e. the width of the central VF must be 20 degrees wide). The US Social Security Administration (SSA) has recently determined that automated perimetry can be used for the latter, with an MD of -22 dB in the better eye corresponding to the VF definition of statutory blindness (a landmark utilised recently by another study) (Heijl et al. 2011).

A question of major interest, however, is how many patients undergoing treatment are progressing at rates that put them in danger of experiencing significant impairment of visual function or legal blindness in their predicted remaining lifetime. This chapter uses more clinical data and cut-offs derived from the research in Chapter 2 of this dissertation and the US SSA to investigate the proportion of patients in clinical practice that are in danger of progressing to levels of meaningful visual impairment and statutory blindness.

The work in this chapter has formed a paper published by *Investigative Ophthalmology and Visual Science* (Saunders et al. 2014). Richard Russell (RR), Jim

Kirwan (JK), Andrew McNaught (AM) and David Crabb (DC) were joint authors of the paper; RR contributed to the data analysis and the work was directed by RR and DC. I prepared the data, performed all the analysis, wrote the paper and produced all the results and figures; RR, DC, JK and AM were all involved in editing and revising the paper. This research was also presented as a read paper at the Royal Statistical Society's Young Statistician's Meeting in London, UK on the 4th July, 2013. The data utilised for this chapter came from Moorfields Eye Hospital, Cheltenham General Hospital Gloucestershire Eye Unit and the Queen Alexandra Hospital in Portsmouth.

3.1. Background

Past studies suggest that many patients newly-diagnosed with glaucoma are not at a high risk of blindness. Studies based on retrospective chart reviews have found that the proportion of patients that progressed to blindness during follow-up ranged from approximately 6% to 13% (Hattenhauer et al. 1998, Kwon et al. 2001, Chen 2003). However, these studies used data collected on manual perimetry, were based on relatively small numbers (all less than 300 patients) and were all carried out more than 10 years ago.

Other estimates for VF loss in predicted lifetime can be extrapolated from more recently conducted prospective studies. The EMGT found that the median progression rate of MD, even in patients without treatment, was slower than -0.5 decibels (dB)/year (Heijl et al. 2009); as an example, a patient with little VF damage at diagnosis (say -2dB) would take 40 years to reach an MD of -22dB at this rate of decay, assuming a linear rate of VF deterioration. Of course, 40 years is likely to exceed most patients' expected lifetimes when it is considered that the onset of glaucoma is usually towards the end of patients' lives. For instance, the mean baseline age of patients in the database used in this thesis is 65 years (excluding patients under 40), although this can vary a large amount. Obviously, there is large variation in rates of VF loss and faster progressing patients are at greatest risk of visual impairment, yet the prevalence of these 'fast progressors' in clinical practice

is unclear. Cohort studies from Canada and California have indicated that less than 5% of patients progress at a rate of -1.5dB/year or worse (Chauhan et al. 2010, Reis et al. 2012, Medeiros et al. 2012). However, results from retrospective studies, which may reflect clinical practice more closely but are less well controlled, are less consistent; results from the New York Progression Study suggest that the prevalence of fast progressors ranges from 9% to over 25% (with the use of different exclusion criteria) (De Moraes et al. 2009b, Prata et al. 2010, Teng et al. 2010) In addition, a recent review of clinical data from Sweden by Heijl and colleagues recorded over 15% of patients progressing at a rate faster than -1.5dB/year (Heijl et al. 2012a).

Taking treatment and monitoring costs into account, it is extremely important that resources are prioritised in favour of those patients that are at greater risk of suffering significant visual disability in their lifetime (Heijl 2013). In this context, information about rate (speed) of VF loss over a period of follow-up is clinically useful. Recent research has emphasised the clinical importance of this approach by recommending frequency and pattern of VF testing required over time in order to establish reasonable estimates of these rates (Chauhan et al. 2008b, European Glaucoma Society 2008, Jansonius 2010, Crabb & Garway-Heath 2012). Once the rate of VF loss has been established, a natural next step is to consider the likelihood of a patient suffering visual disability within their expected lifetime. Incorporating estimates of life expectancy adjusted for age (residual life expectancy) in glaucoma care was suggested by Wesselink et al. (Wesselink et al. 2011), but before now has not been implemented. Therefore, this study attempted to combine information about speed of VF loss and residual life expectancy to estimate the proportion of patients under clinical care in glaucoma clinics that progress at a rate quick enough to result in serious visual impairment in their expected lifetime.

3.2 Materials and Methods

Three separate Medisoft (Medisoft Ltd., Leeds, UK) VF databases from glaucoma clinics in Moorfields Eye Hospital in London, Cheltenham General Hospital

Gloucestershire Eye Unit and Queen Alexandra Hospital in Portsmouth were available for this retrospective study. The databases contained a total of 402,357 anonymised VFs from 75,857 patients recorded between 1989 and 2012. The study adhered to the tenets of the Declaration of Helsinki, was approved by a research ethics committee of City University London and all anonymised data were transferred to a secure computer database. Only VFs recorded on the HFA with the 24-2 test pattern, a Goldmann size III stimulus and the SITA (SITA Standard or SITA Fast) testing algorithm were included in the study. Unreliable VFs, defined here to have a FP or FN score of 33% or more, or, a FL score of 20% or more were excluded; 29.6% of the SITA Standard and Fast VF records were excluded from the study owing to these reliability criteria.

This study centred on VF progression, and, to be included in the study, one of each patient's eyes had to have a VF series that was at least 3 years long, with at least 5 VFs after discounting the first VF in order to attempt to obviate learning effects (improvements in results through the patient becoming more practiced at taking the tests) (Wild et al. 1989, Heijl et al. 1989, Heijl & Bengtsson 1996). The first recorded MD, having excluded the initial VF for learning effects, is subsequently referred to as the baseline MD, whilst the last recorded MD in the follow-up is denoted as the final MD. The study only considered VF data and no other clinical information. Therefore, it was not possible to confirm whether individuals in the database were *clinically* diagnosed with glaucoma or were glaucoma suspects. Thus, the baseline VF of each patient had to have an HFA MD or HFA pattern standard deviation value outside the established HFA 95% normal limits in at least one eye. Given that individuals had measurable VF damage at baseline and were followed regularly for at least 3 years in these glaucoma clinics, it is reasonable to assume that the vast majority in this study had glaucoma. Patients under 35 years of age were not included in the study.

3.2.1 Extrapolating Visual Field Status at Patient End of Expected Lifetime

Rates of MD loss were calculated in decibels per year (dB/y) using ordinary least

squares regression. It was assumed that the rate of change in MD would remain constant for the remainder of a patient's expected lifetime. Therefore, the MD at expected death was calculated as the observed rate of loss multiplied by the patient's residual life expectancy (**Figure 3.1**). It is important to use residual life expectancies to take into account patient age rather than use population age statistics, as the probability of dying at a particular age will not remain constant across all ages. For instance, the probability of a 10 year old living to 78 is very different to an individual that has already lived to 76. Median life expectancies, based upon age and gender, were collected from the UK ONS (Office of National Statistics 2011); these were derived using the latest available English census data (2001) and survey data to estimate expectancies valid for the 2008-2010 period.

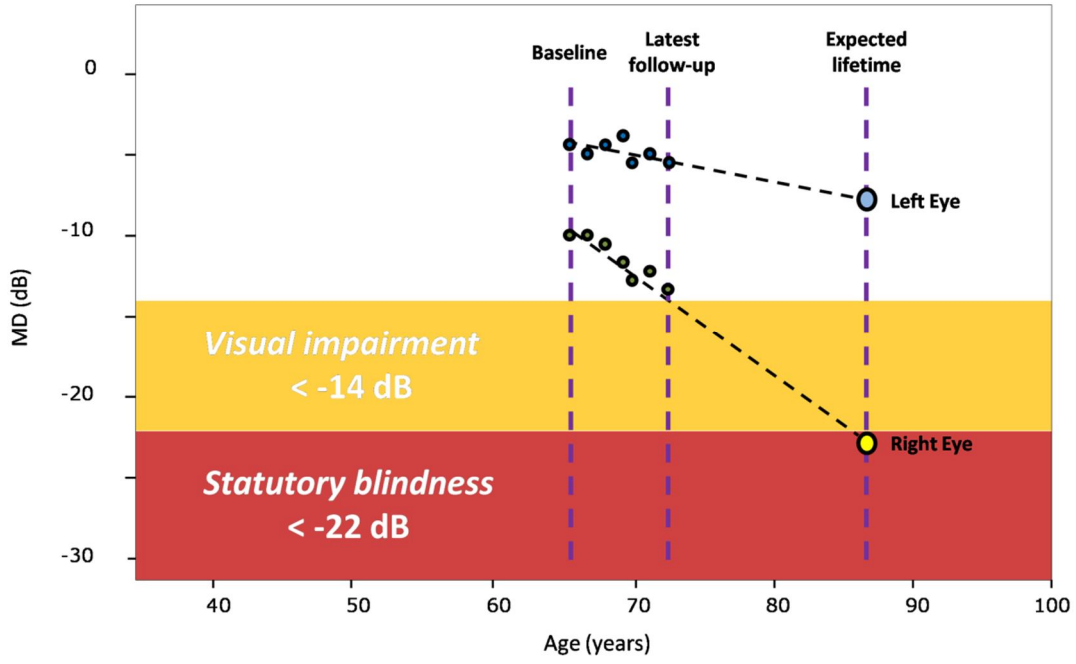

Figure 3.1 - A schematic illustrating the analysis conducted in this study. Visual field (VF) series from the left and right eyes of a patient were used to estimate a linear rate of loss in each eye (dB/y). The patient's median life expectancy was obtained from the UK Office of National Statistics (Office of National Statistics 2011), and was used to predict the mean deviation (MD) of each eye at expected time of death. In this illustration, the right eye was anticipated to progress into the statutory blindness stage by the end of the patient's life. However, given that the left eye is progressing less quickly and has less VF damage at the outset, this patient would be unlikely to experience severe visual disability in their lifetime.

3.2.2 Characterising the Expected Visual Function of each Patient

When assessing visual disability in a patient's lifetime, it is necessary to consider the VF loss in both eyes; a damaged VF in one eye will not necessarily impact on the overall QoL of a patient if the other eye remains healthy (Jampel, Friedman et al. 2002). Methods for classifying overall VF damage into 'moderate' or 'severe' abound (Hodapp et al. 1993, Brusini & Filacorda 2006, Susanna Jr. & Vessani 2009), but few offer thresholds that have an evidence base underpinning their link to visual impairment or offer any practical meaning to a patient. Nevertheless, results from Chapter 2 suggested that a patient with an MD less than -14dB in their "better" eye (the eye with the least damaged VF) is highly unlikely to satisfy the vision component of legal fitness to drive in the UK. Hence, this threshold was employed as a benchmark for "visual impairment" in the current study. A second MD benchmark for "statutory blindness" equal to -22 dB, as recommended by the US SSA (US Social Security Administration 2011), was also employed. To be classified as having visual impairment or statutory blindness both patient eyes had to reach the necessary levels of VF loss. If patients had less than 3 VFs performed in one of their eyes then their baseline MDs were recorded and the eyes were either deemed to be stable (i.e. progressing at a rate of 0 dB/year) or progressing (at a rate of -1.5 dB/year) representing a 'best' and 'worst' case scenario respectively. In the event of no recorded VF for the other eye, the eye was recorded as either visually healthy for the 'best case scenario' or blind in the 'worst case scenario'. The open-source statistical environment R (R Development Core Team 2014) was used for all statistical analyses.

3.3 Results

The characteristics of the study sample of 3790 patients are given in **Table 3.1**. **Figure 3.2 (A) and (B)** show the distribution of patient eye follow up times, patient residual life expectancies and progression rates in all eyes, respectively. It is apparent from **Figure 3.2 (B)** that the vast majority of eyes progressed at a rate between ± 0.5dB/year (74.0% - 95% binomial CI: 73.0 to 75.0%). A small proportion of patient eyes progressed at a rate worse than -1dB/year (7.5% - CI: 6.9 to 8.2%) and only 3.0% (CI: 2.7 to 3.4%) of eyes progressed at faster than -1.5dB/year. It is worth noting that a considerable number of eyes recorded positive MD rates (33.3% - CI: 32.2 to 34.4%).

Table 3.1 - The demographics of patients analysed in the study. IQR denotes the interquartile range, whilst the better eye corresponds to the eye with the better mean deviation (MD) at the baseline VF examination

Measure	Patients with series in both eyes (n=3359)	Patients with a series in one eye only (n=431)
Number of patients with glaucoma (VF defects) in both eyes (%)	2212 (65.9)	N/A
Number of female patients (%)	1684 (50.1)	228 (52.9)
Number of patients from London (%)	3124 (93.0)	413 (95.8)
Number of patients from Cheltenham (%)	183 (5.4)	10 (2.3)
Number of patients from Portsmouth (%)	52 (1.5)	8 (1.9)
Median number of VFs recorded (IQR)	6 (5 to 8)	6 (5 to 8)
Median follow-up time (IQR)	7.1 years (5.2 to 9.1)	7.2 years (5.5 to 8.9)
Median Baseline Age (IQR)	65 years (56 to 72)	66 years (57 to 75)
Median Final Age (IQR)	71 years (62 to 78)	72 years (63 to 80)
Median Residual Life Expectancy from final age (IQR)	16 years (11 to 22)	14 years (9 to 21)
Median Baseline MD in better eye (IQR)	-2.6dB (-5.2 to -1.1)	-7.0dB (-11.9 to -3.7)
Median Baseline MD in worse eye (IQR)	-6.9dB (-12.5 to -3.8)	
Median Final MD in better eye (IQR)	-3.4dB (-6.8 to -1.3)	-8.4dB (-14.6 to -4.7)
Median Final MD in worse eye (IQR)	-8.7dB (-14.8 to -4.6)	
Median Rates of Loss in better eye (IQR)	-0.12dB/year (-0.38 to 0.07)	-0.19dB/year (-0.54 to 0.05)
Median Rates of Loss in worse eye (IQR)	-0.15dB/year (-0.46 to 0.08)	

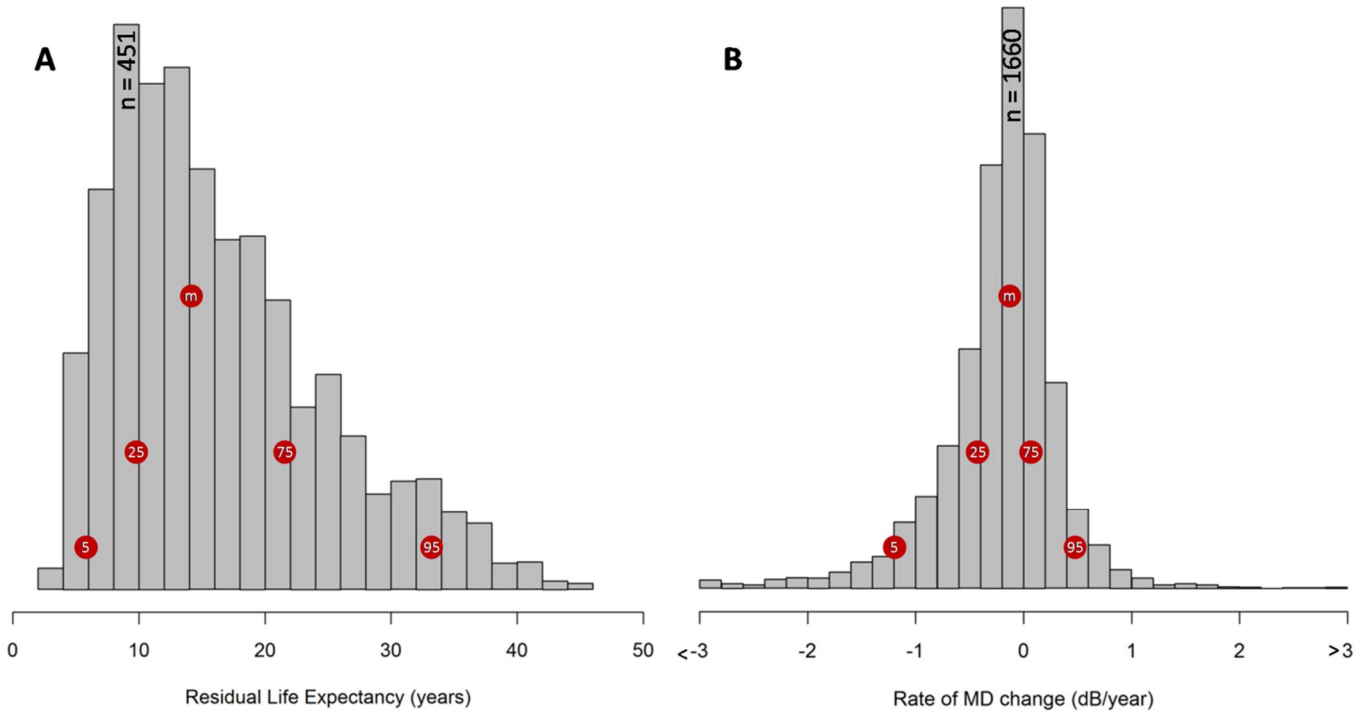

Figure 3.2 (A) Distribution of residual life expectancies for all 3790 patients included in the study and (B) the rate of progression of Mean Deviation (decibels per year) from all 7149 eyes. The distribution of life expectancies is positively skewed as a result of the increased prevalence of glaucoma in older patients. The red circles indicate the median (m) and other quantiles.

Of the 3359 patients with a VF series from both eyes (**Table 3.2** and **Figure 3.3**), only 5.2% progressed to statutory blindness (both eyes progressing to an MD worse than -22dB) with a further 10.4% progressing to visual impairment (both eyes progressing to an MD level of worse than -14dB) in their expected residual lifetime. The 'best case scenario' produced similar results to those just considering eyes with two series, but under the 'worst case scenario', the number of patients at risk of statutory blindness increased to 7.1%, and a further 11.5% were at risk of visual impairment (**Table 3.2**). Interestingly, almost half of the patients with both eyes followed had at least one eye with a positive rate of change (49.0% - CI: 47.3 – 50.7).

Table 3.2 - The proportion of patients likely to suffer visual field impairment in the course of their lifetime

Visual impairment at death	% No impairment (95% CI*)	% Visual impairment (95% CI*)	% Statutory Blindness (95% CI*)
Including patients with a series for each eye only (n=3359)	84.4 (83.2 – 85.6)	10.4 (9.4 – 11.4)	5.2 (4.5 – 6.0)
All patients best-case scenario (n=3790)	84.9 (83.7 – 86.1)	10.0 (9.0 – 11.0)	5.1 (4.3 – 5.8)
All patients worst case scenario (n=3790)	81.5 (80.2 – 82.8)	11.5 (10.4 – 12.5)	7.1 (6.2 – 7.9)

*95% confidence intervals (CI) were calculated using the normal approximation of a binomial distribution

When just patients with series in both eyes tested were considered, 159 of the 175 patients (90.9% - CI: 86.6 to 95.1%) who reached statutory blindness had an MD worse than -6dB in at least one eye at baseline; this MD level is equivalent to what

is considered to be at least a “moderate defect” for one criterion of the Hodapp-Parrish-Anderson index (Hodapp et al. 1993). Patients that were predicted to progress to statutory blindness were around 70% more likely to have moderate damage (MD worse than -6dB) in at least one eye at baseline than patients not predicted to progress to this stage (+ Likelihood Ratio: 1.7; 95% CI: 1.6 to 1.8). Put differently, only 1.1% (CI: 0.6 to 1.6%) of the patients that were likely diagnosed with early VF defects, with an MD better than -6dB in both eyes (44% of the study population), progressed to statutory blindness. Strikingly, almost 60% (CI: 52.0 – 66.4%) of patients progressing to statutory blindness had one eye with an MD already worse than -14dB in at least one eye at baseline.

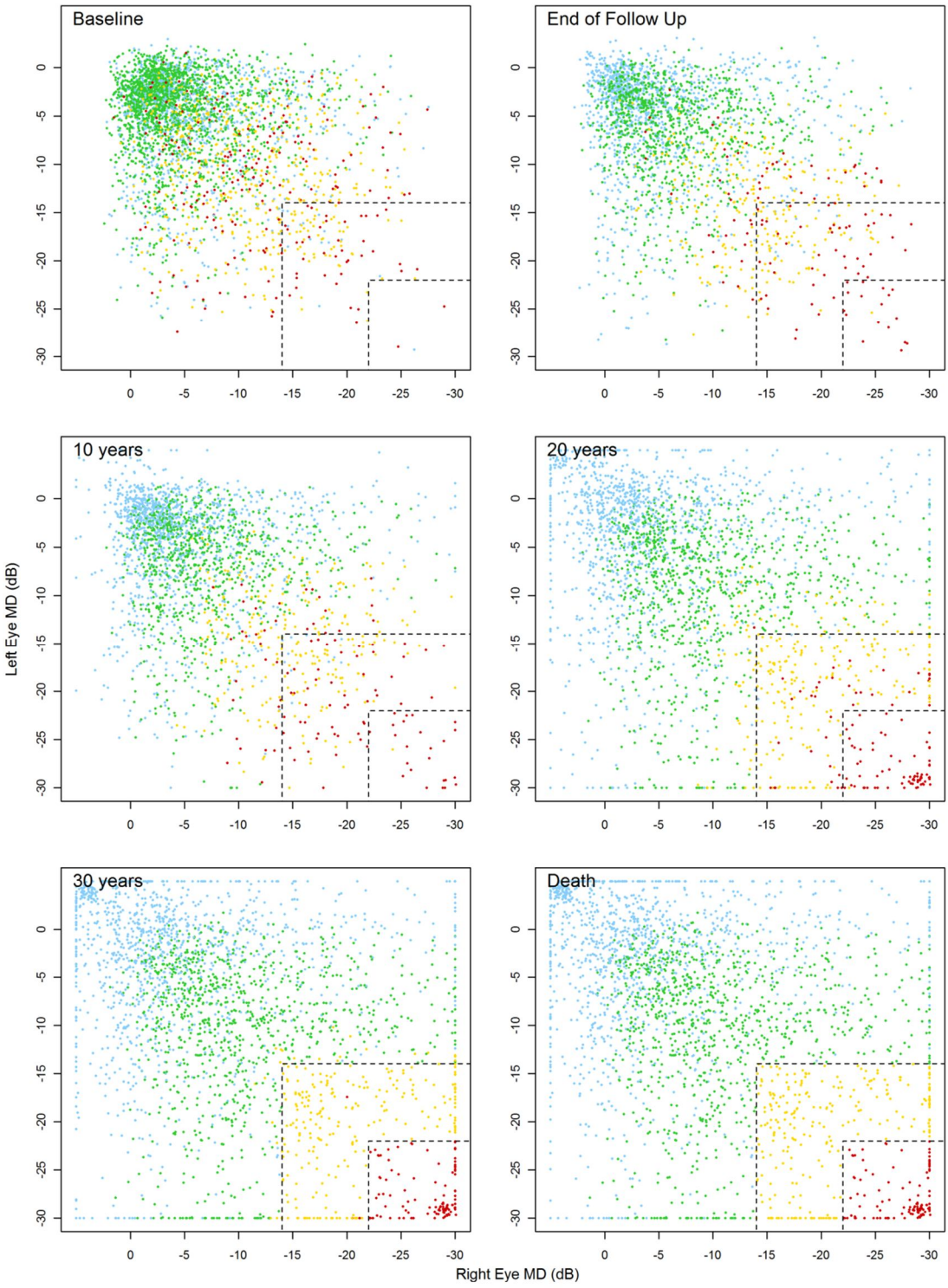

Figure 3.3 - A series of scatterplots showing Mean Deviation (MD) in vertical (Y-axis) and horizontal (X-axis) eyes at baseline, at the end of follow-up and, through extrapolating current rates of MD deterioration, after 10, 20 and 30 years follow-up and at the end of expected lifetime. Both eyes in the plot had to fulfil the original inclusion criteria. The patients are coloured according to their visual disability status at expected time of death. Blue represents a patient where at least one of the eyes has a positive slope over time, green represents progression, but no significant impairment by the end of the patient's lifetime, yellow represents degradation to visual impairment (-14dB or worse in both eyes), while red corresponds to statutory blindness in both eyes (below -22dB).

3.4 Discussion

This retrospective study of a very large number of VFs collected in different clinics over a 23 year period has provided a number of interesting findings. The modelling indicated that a small proportion of patients under clinical care in glaucoma clinics in the UK were estimated to be at risk of progressing to a level of statutory blindness in both eyes during their predicted residual lifetime. The proportion of patients predicted to be at risk of progressing to statutory blindness in both eyes within the study was 5.2%, though this figure may be as high as 7.1% ('worst case scenario') depending on the reasons behind only testing single eyes in some patients. These results, from the perspective of the burden of diagnosed glaucoma, seem more optimistic than those of previous studies. For example, Kwon et al. predicted from their study that the number of patients becoming legally blind over a follow-up of 22 years could be as high as 19% (Kwon et al. 2001), whilst others have also predicted higher proportions (Hattenhauer et al. 1998, Chen 2003). The different methodologies used in these studies conducted more than a decade ago are likely to explain the different results. For example, the previous studies did not use a modelling approach, were based on far fewer patients, used different definitions for legal blindness and because they used 'retrospective chart review' were very likely subject to selection bias. At the same time, it is tempting to explain the differences with the idea that modern therapies are improving patient prognosis in glaucoma. Despite small numbers reaching statutory blindness it should be noted that a significant minority of patients (around one in six patients) in this study were predicted to develop VF loss that could affect their quality of life; for example, a level of impairment that would likely result in loss of a driving license in the United Kingdom as shown in Chapter 2 (Saunders et al. 2012a).

Interestingly the very wide distribution of rates of VF loss shown in **Figure 3.2 (B)** is reminiscent of similar results shown in controlled prospective studies (Heijl et al. 2009, Chauhan et al. 2010). However, the proportion of eyes that are very rapidly progressing appears substantially smaller than those of many other retrospective studies. The 3% of eyes highlighted in this study as progressing at faster than -

1.5dB/year was in contrast to the figures from the recent findings from Heijl et al.'s study in Sweden that estimated that 15% progressed at a rate faster than -1.5dB/year (Heijl et al. 2012a) and the New York Progression Study, which concurred that this proportion was in excess of 9% (De Moraes et al. 2009b, Prata et al. 2010, Teng et al. 2010). There are several possible reasons behind this difference. First, a sizeable proportion of pseudoexfoliation glaucoma (associated with faster disease deterioration (Heijl et al. 2009)) was present in the Heijl study, not as commonly seen in the UK. Another possible cause is that the patients in these studies were diagnosed with more advanced glaucoma, although it is not really known whether those with advanced defects progress more quickly or whether they have reached a stage of more advanced impairment because they presented later. It is perhaps worth remembering that sensitivities below 20dB become highly variable, so, as MDs decrease, they are likely to be subject to increasing noise also. This in turn could contribute to estimates of rate of loss more variable too, which may be a factor behind some faster rates of loss exhibited. Nonetheless, this study's estimates of the 'fast progressor' prevalence was more akin to those results from controlled clinical cohort studies (Chauhan et al. 2010, Medeiros et al. 2012, Reis et al. 2012).

Remarkably, about half of all the patients sampled experienced an "improved" MD in at least one eye during their follow-up, as can be seen in **Figure 3.3**. These positive slopes can be explained by a combination of high variability associated with VF measurements (Heijl et al. 1987, Turpin & McKendrick 2011, Russell et al. 2012a) and learning effects, which can persist over 10 or more tests in some individuals (Heijl et al. 1989, Heijl & Bengtsson 1996). The study attempted to control for the latter by adopting the common practice of removing the patient's first recorded VF. However, there evidently remains a substantial difficulty in measuring rates of MD change.

About 90% of those patients predicted to be at risk of statutory blindness in their residual expected life time already had noteworthy VF damage (MD worse than -6dB) in at least one eye at baseline (see **Figure 3.3**). Most of these patients had advanced impairment (60% worse than -14dB in at least one eye) at baseline.

These statistics strongly suggest that a major contributory factor for the risk of future visual impairment, or statutory blindness, from glaucoma is late presentation of the disease. Other studies support this important notion (Grant & Burke 1982, Oliver et al. 2002) and some have explored it in more detail, highlighting the real threat to blindness associated with the late detection of the disease (Sinclair et al. 2004, Henson & Thampy 2005, Kotecha et al. 2012a). Indeed, given this was a study necessarily limited only to patients that were under clinical care, these results suggest that it is the many undiagnosed glaucoma sufferers, who are at the greatest risk of blindness. This finding raises an interesting debate about how best to balance the use of VF resources in primary and secondary care, especially since it was estimated that only a minority of diagnosed patients in clinical care are in danger of being severely impacted by their condition during their lifetime. Specifically, it suggests that more resources should be directed towards detecting and case finding glaucoma. Moreover, whilst glaucoma is a chronic disease, the data here highlights that those affected are, of course, typically elderly and have low residual life expectancy; the results from this study should reinforce the need for clinicians to consider life expectancy in their clinical management of the disease (Wesselink et al. 2011)

3.4.1 Retrospective Data Analysis

Results from clinical trials and prospective studies primarily inform clinical practice and decisions about health service delivery. Still, retrospective analysis of very large volumes of data collected from the everyday clinical milieu over long periods of time can provide interesting material and information to develop new hypotheses, as this report shows. It is already known that volunteers for prospective studies in glaucoma have better adherence to prescribed therapy than those in routine medical care (Henson & Shambhu 2006), so prospective studies and trials may even misrepresent the routine clinical situation. However, any retrospective study, including this one, will have issues with missing or incomplete data, although, in this case, this was largely offset by the sheer volume of data; a significantly larger number of VFs were utilised compared to other past retrospective studies. One consequent limitation of the retrospective study was

that a small proportion of the patients studied only had a complete series in one eye (431 out of 3790 patients included in the study). Unfortunately, the reasons behind the other eye not being followed were disparate: the eye may have either had extensive damage that rendered 24-2 VF tests uninformative (the patient may have been switched to 10-2 testing for example), or else it may have been healthy and not tested frequently enough to meet the inclusion criteria. It was attempted to cater for both of these possibilities by providing both the 'worst case scenario' and 'best case scenario' results. Another issue associated with this retrospective analysis concerns the fact that full patient records were unavailable or not considered. As a result, analyses were just based on age, gender and VF data. Some of the faster progressing patients in this sample may have therefore had rapidly lowering MDs as a result of concomitant age-related eye disease, principally cataract. It is possible not every patient in the study sample may have had glaucoma, but this possibility was minimised through constraining subjects to those being followed over at least 3 years; the majority of subjects that were monitored at specialist glaucoma clinics over at least 3 years are likely to have glaucoma, after all. Furthermore, these findings must be tempered by the possibility that the baseline fields of the patients may not have been their first VF assessment; for instance, patients may have been transferred from a different clinical centre.

3.4.2 Strengths and Limitations of the Study

One strength of this study is that the thresholds chosen for visual impairment and statutory blindness have some evidence-based justification attributed to them. Reaching levels of MD worse than -14dB in both eyes has been shown to correspond with being legally unfit to drive in the UK, amongst difficulties carrying out various other visual tasks (Kooijman et al. 2004, Kotecha et al. 2009, Smith et al. 2011, Black et al. 2011, Glen et al. 2012, Tabrett & Lathan 2012). Furthermore, -22dB in both eyes is the point at which one qualifies for statutory blindness in the US, so represents a significant milestone for patients. However, in spite of the fact that measured sensitivities are weighted towards fixation when calculating MDs, it is, of course, possible to have preserved visual acuity under these conditions; people who are diagnosed with legal blindness can still have some useable vision.

In addition, it is important to emphasise that using the MD to define visual disability does not appreciate the spatial distribution of VF damage, which is important in a patient's visual function, and ability to carry out different tasks (Kooijman et al. 2004, Kotecha et al. 2009, Black et al. 2011, Tabrett & Lathan 2012). For example, VF loss close to fixation is particularly important and eyes with this damage should be treated more aggressively, especially due to the fact that the likelihood of further damage in the central VF is higher (Membrey et al. 2000).

An assumption of this analysis involves the use of a linear rate of progression of MD over time. This may not reflect the true nature of glaucomatous deterioration given that there is some evidence to show patients tend to progress more quickly at older ages, although it is unknown whether this is a result of older age or more advanced VF deterioration (Heijl et al. 2012a). Some researchers have also advocated the use of exponential models for monitoring VF decay for individual test sensitivities in particular (Caprioli et al. 2011; Azarbod et al. 2012; Pathak et al. 2013). Nevertheless, linear regression of MDs is commonly utilised in clinical practice; the Glaucoma Progression Analysis software in the HFA, for instance, presents this as "one method of Tracking Rate of Progression" (Carl Zeiss Meditec 2003). Furthermore, studies suggest that linear rates of progression for summary measures are adequate (Bengtsson et al. 2009). In addition, previous work has shown that a linear model of VF progression tends to provide more robust estimates of future measurements than more complex models (McNaught et al. 1995, Bryan et al. 2013). This demonstrates a simple statistical principle that, although more complex models tend to provide better fits of existing data, linear models tend to be more useful at predicting future change. However, it is important to be aware that this MD regression does not imply a constant rate of sensitivity loss – a loss of 1dB implies much more damage going from -5dB to -6dB than from -25 to -26dB, as a result of the logarithmic scaling utilised for the measurement. It is further noteworthy that the 'future' forecasts based upon current linear rates of VF loss may make estimates of future prognosis in the patients studied overly pessimistic, as treatment is usually intensified if a patient is in danger of progressing to visual disability (Heijl 2013). On the other hand, the

modelling takes no account of concomitant eye disease which ultimately might precipitate levels of lifetime visual disability that are worse than those shown in the results. A technical limitation of the calculations is that life expectancies utilised were periodic and so the study assumes that the probability of dying at a given age will remain constant over time. It was also necessary to assume that mortality rates are independent of glaucoma as a condition, although this seems reasonable given results from other studies that have specifically looked at this (Grodum et al. 2004). Mortality rates could be affected by other morbidities which may be more common in individuals with glaucoma; it is further possible that fast progression rates may be symptomatic of poor general health or access to medical care, which may in turn affect life expectancy. However, more research is required to fully understand how life expectancy corresponds with rates of VF loss before this can be taken into consideration.

It is important to reflect on the generalisability of results. Firstly, the sample was only composed of patients from hospital care in the UK, and it should be further noted that the number of patients from hospitals in Portsmouth and Cheltenham were considerably fewer than those from Moorfields Eye Hospital in London. Of course, there is no guarantee that patients were treated equivalently across hospitals and thus these results mostly reflect observations in Moorfields Eye Hospital. Results are not directly applicable to countries with different demographics and different health care systems. On the other hand, London is a cosmopolitan city; the 2011 Census reported that more than one quarter of the population of London did not identify themselves as British (Office of National Statistics 2012), so it could be postulated that there is a fair amount of diversity amongst the patients, although the lack of clinical data makes it impossible to determine this for certain. Another issue with the study, particularly when focussing on the finding that worse VF loss at baseline is associated with a higher risk of blindness, is that it is assumed that various factors that may have an effect on end state of disease, such as type of glaucoma and race of patient, were relatively uniform throughout the distribution. Unfortunately, it was not possible

to control for these variables, but it is assumed, given the quantity of the data, that the findings overall represent a reasonable population estimate.

3.4.3 Conclusions

In conclusion, the main result from this modelling exercise suggests that most glaucoma patients under clinical care are not in danger of suffering significant visual disability in their lifetime. This report certainly indicates that the great majority of patients that are followed in glaucoma clinics in the UK have stable VFs. There is, however, enormous variability in rates of VF loss and also in levels of VF damage at presentation. Patients that are in danger of significant VF impairment in their lifetime generally present with more severe VF damage and this may indicate that more resources should be concentrated towards detecting disease before it progresses beyond early damage. It is hoped that these results can inform the design of better health service delivery and suggest studies that should investigate improved allocation of VF testing resources (Heijl 2013). Furthermore, the results from this study illuminate very clearly the importance of reliably detecting significant VF defects, and other features of glaucoma, in primary care (Henson & Thampy 2005).

No matter how quickly glaucoma is diagnosed, however, it is important to establish rates of loss quickly as there is clearly much variability in the rates of loss and risk of blindness from patient-to-patient and it is important to give higher risk patients prompt intervention. However, this takes time, mainly due to variability in measurements. This can potentially be reduced by increasing the number of tests used to establish thresholds (thereby increasing the certainty of a measured threshold), but also by reducing the duration of the test, which can prevent patient fatigue. Essentially, VF testing algorithms are a compromise between these two considerations (Turpin & McKendrick 2011). SITA Standard is a current gold standard for measuring threshold in automated perimetry, but SITA Fast is a quicker test to complete (Nordmann et al. 1998, Wild et al. 1999, Pierre-Filho et al. 2006), although it has been suggested that it may be less precise (Artes et al. 2002). The next chapter will therefore look at the differences in test-retest variability between

SITA Standard and SITA Fast perimetric testing algorithms and attempts to establish whether any difference between the tests is clinically meaningful in the context of time to detect progression in clinical practice.

Chapter Four: Comparing the relationship between variability and sensitivity in SITA Standard and SITA Fast visual fields

In the previous chapter it was established that a low but significant proportion of patients in clinical practice progress at a rate that could lead to visual impairment or blindness. In other words, the majority of patients under clinical care are not likely to go blind in their predicted remaining lifetime (Saunders et al. 2014). It is important to be able to differentiate between stable and dangerously progressing patients in order to ensure appropriate treatment can be given to prevent sight loss. Measuring progression of the VF plays an important role in finding out which groups require treatment intensification. However, VFs often yield variable measurements. Moreover, measurement variability increases with declining VF sensitivity (Henson et al. 2000, Artes, Iwase et al. 2002, Russell et al. 2012a). As a result of this variability, detecting VF progression is far from straight-forward, with significant potential for detecting false change or failing to detect definite change.

Visual field testing algorithms implemented in SAP ought to generate measurements with precision and accuracy that are sufficient for monitoring progression effectively. Accuracy and precision are terms that are often confused: an accurate test is one that produces results with as little bias from true measurements as possible, whereas a precise test is one with high repeatability in the test results. The precision of a VF test can be optimised by repeated or extended testing, but this is offset by the demand for quick examination in a clinical environment and a need to reduce the effect of fatigue on test performance (Henson & Emuh 2010).

SITA Standard was developed for the HFA in the 1990s and allows VF testing to be performed in about half the time of that taken by older 'Full Threshold' strategies (Nordmann, Brion et al. 1998, Bengtsson & Heijl 1998a, Wild et al. 1999) with no significant decrease in precision (Bengtsson & Heijl 1998a, Artes et al. 2002). SITA Standard uses prior information based on the sensitivities of the surrounding points

and stimulus sequences are interrupted when the measurement error of the test points is small compared to the ERF (Bengtsson et al. 1997a, Heijl et al. 2012b). In fact, SITA standard has become a clinical standard for acquiring VF measurements. SITA Fast provides an even quicker test time often taking 5 minutes or less on average to administer, through presenting starting stimuli closer to expected thresholds and because stimulus staircases are interrupted at an earlier stage by increasing the ERF cut-off, thereby accepting lower accuracy of test results (Nordmann et al. 1998, Bengtsson & Heijl 1998b, Wild et al. 1999, Pierre-Filho et al. 2006). The shorter duration of the SITA Fast test means that it is an appealing method from a practical and clinical point of view. Yet there remains uncertainty about the precision of measurements from SITA Fast relative to SITA Standard and how this impacts on time to detect VF progression. Choice of testing algorithm therefore presents a dilemma to the clinician.

The main aim of this chapter is thus to provide some evidence for deciding which test should be used in clinical practice. This is carried out by clarifying the difference in precision between SITA Standard and SITA Fast across the full range of VF sensitivities using the large volumes of patient data from clinical practice in testing centres across the UK. The clinical impact of choice of VF test algorithm is then estimated by considering average time to detect progression using computer simulation.

The work in this chapter has formed a paper that has been published in the *Journal of the American Medical Association Ophthalmology* (Saunders et al. 2015). My co-authors for this manuscript were Richard Russell (RR) and David Crabb (DC). RR contributed to the data analysis and the work was directed by RR and DC. I prepared the data, performed all the analysis, wrote the paper and produced all results and figures; RR and DC edited and revised the paper. This research was further presented as a read paper at two international conferences, including the North American Perimetry Society Meeting in Chapel Hill, NC, USA on the 3rd October 2013 and the Association for Research in Vision and Ophthalmology Annual Meeting in Orlando, FL, USA on 6th May 2014.

4.1 Materials and Methods

This study analysed 473,252 anonymised VFs from 88,954 patients from databases at Moorfields Eye Hospital glaucoma clinic in London (320,334 VFs), Cheltenham General Hospital Gloucestershire Eye Unit (50,144 VFs), Queen Alexandra Hospital in Portsmouth (31,879 VFs) and the Calderdale and Huddersfield NHS Foundation Trust (70,955 VFs). The study adhered to the Declaration of Helsinki, was approved by a research governance committee of City University London and all anonymised data were transferred to a secure database. Only VFs from the HFA using Goldmann size III stimuli with the 24-2 test pattern and either SITA Standard or Fast testing algorithms were included in the study. Eyes were excluded if they had fewer than five VF examinations; if both eyes fulfilled these criteria for an individual patient then one eye was selected at random. In addition, patients under 35 years old were also excluded from the study. The first VF was removed in order to attempt to account for perimetric learning effects (Wild et al. 1989, Heijl et al. 1989, Heijl & Bengtsson 1996) leaving a total of 86,793 VFs from 13,778 patients for analysis; 66,974 SITA Standard VFs from 10,124 patients and 19,819 SITA Fast VFs from 3,654 patients.

4.1.1 Determining the precision of SITA Standard and SITA Fast

The precision of SITA Standard and SITA Fast was assessed by fitting ordinary least squares regression to each VF test location's sensitivity series (i.e. PLR). Fitted values were obtained from the PLR model for each test location (excluding the blind spot test locations) and raw residuals were then calculated through subtracting the calculated fitted values from the observed sensitivities at each time point in the individual series. The raw residuals were pooled and binned according to fitted sensitivity (rounded to the nearest decibel). Precision was defined to be higher where raw residuals were smaller. This approach, described in more detail by Russell et al. (Russell et al. 2012a), was carried out separately for eyes tested using SITA Standard and Fast so that the precision of the two methods could be compared.

4.1.2 Evaluating the impact of differences in precision on time to detect progression in clinical practice

In order to illustrate the clinical relevance of any additional variability associated with either of the two algorithms, computer simulations were carried out to determine how soon VF progression could be diagnosed. The methodology utilised to simulate VF progression is reported in more detail elsewhere (Russell et al. 2013). Ten thousand pointwise VF series were simulated to deteriorate at three different rates, or speeds, of VF loss over time (-0.5, -1 and -2dB per year) from three separate starting sensitivities (30dB, 20dB and 10dB). Variability was derived from the distributions of PLR residuals extracted for each algorithm and input into the computer simulation. The time taken to detect deterioration at the $p<0.01$ significance level with two tests per year were recorded and median detection times were compared between SITA Standard and Fast. All statistical analyses were carried out in the open-source programming language, R (R Development Core Team 2014).

4.2 Results

In total, 716,456 PLR models were fitted across all 52 test locations (blind-spots locations excluded) of the 13,778 eyes studied, including 526,448 regression lines for SITA Standard test series and 190,008 regression lines for SITA Fast. As a result, 4,508,036 residuals were generated. Characteristics of the study sample are given in **Table 4.1**. Baseline MDs indicated that the majority of patients included in the study did not have severe VF loss, although 1281 SITA Standard eyes (12.7%) and 321 SITA Fast eyes (8.8%) had baseline MDs of -12dB or worse. Baseline MDs for SITA Fast tended to be better than those for SITA Standard. MD and pointwise progression rates (dB/year) in the two groups were very similar. Interestingly, people followed using SITA Fast tended to be older than those followed using SITA Standard.

Table 4.1 – Characteristics of Study Sample

Measurement	SITA Standard (10,124 eyes)	SITA Fast (3,654 eyes)
	Median (Interquartile Range)	
Number of VF tests	**6** (4 to 8)	**5** (4 to 6)
Baseline VF sensitivity (dB)	**27** (23 to 30)	**28** (24 to 30)
Baseline Mean Deviation (dB)	**-3.23** (-7.37 to -1.03)	**-2.20** (-5.42 to -0.38)
Pointwise progression rate (dB/year)	**-0.13** (-0.55 to 0.17)	**-0.16** (-0.62 to 0.17)
MD progression rate (dB/year)	**-0.11** (-0.43 to 0.12)	**-0.14** (-0.49 to 0.09)
Baseline age (years)	**64** (53 to 72)	**70** (61 to 78)
Follow-up period (years)	**6.0** (4.0 to 8.5)	**5.1** (3.2 to 7.3)

4.2.1 The relative precision of SITA Standard and SITA Fast

Distributions of residuals for each fitted-sensitivity level are shown in **Figure 4.1**. As observed in previous studies, the distributions of residuals varied according to sensitivity (Artes et al. 2002, Russell et al. 2012a). More noteworthy were similarities between SITA Standard and Fast results across all sensitivities. Plotting the standard deviation for each fitted sensitivity (**Figure 4.2**) reveals little difference in precision between the two test algorithms until sensitivity has declined below around 20dB. In fact, the precision of SITA Fast appears to even be slightly improved over SITA Standard at higher sensitivities (i.e. the spread of residuals is smaller). The difference in precision between the two tests peaks at just under 15dB, where the test variability is also at its highest in both tests. Even then the magnitude of the difference is small: for example, at 14dB the difference in standard deviations is 0.51dB and this represents a difference of only about 8%.

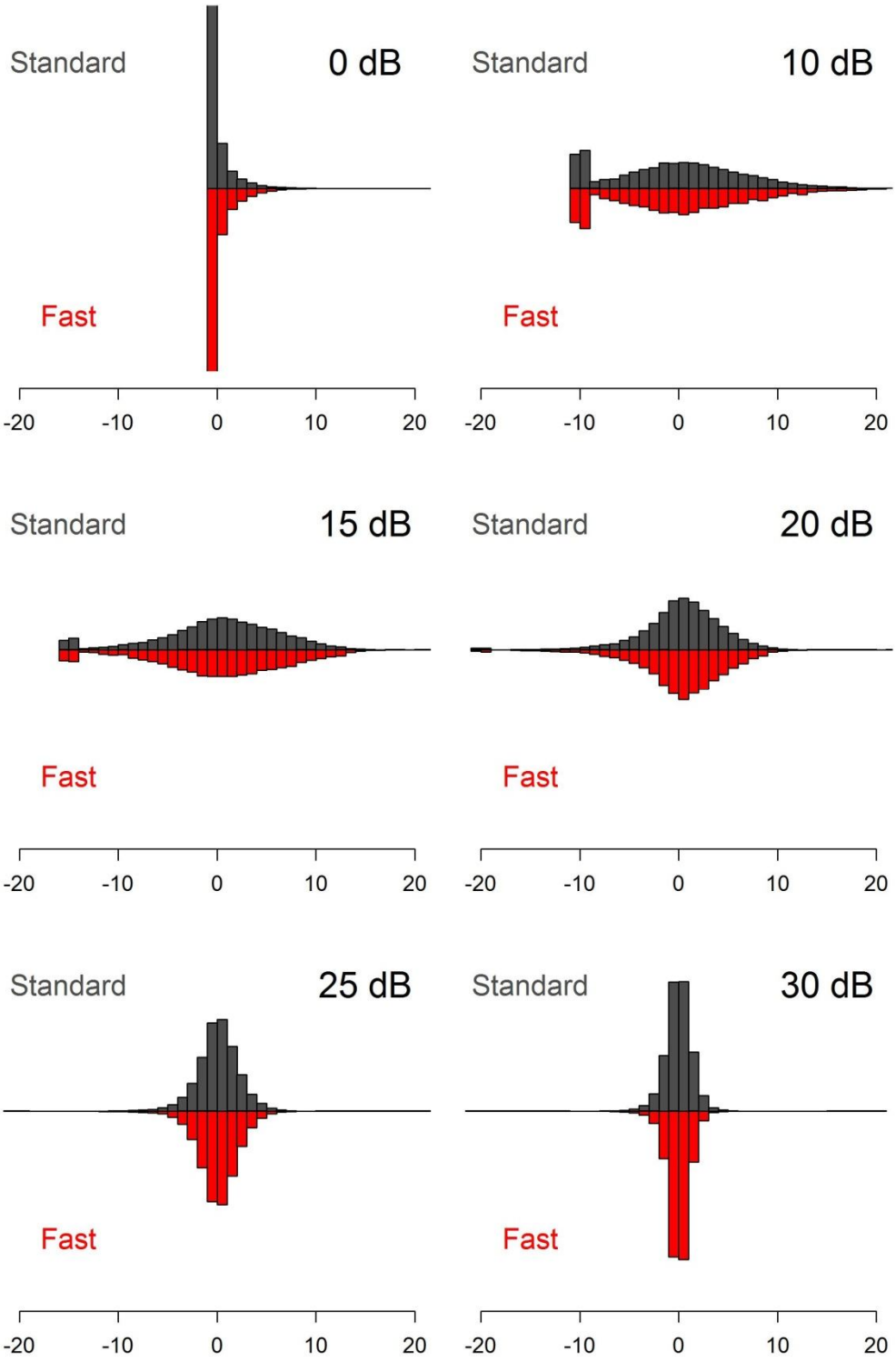

Figure 4.1 - Back-to-back histograms showing distributions of the frequency density of raw residuals generated through linear modelling of each test location at rounded fitted values of 0, 10, 15, 20 and 30dB for SITA Standard (grey) and SITA Fast (red).

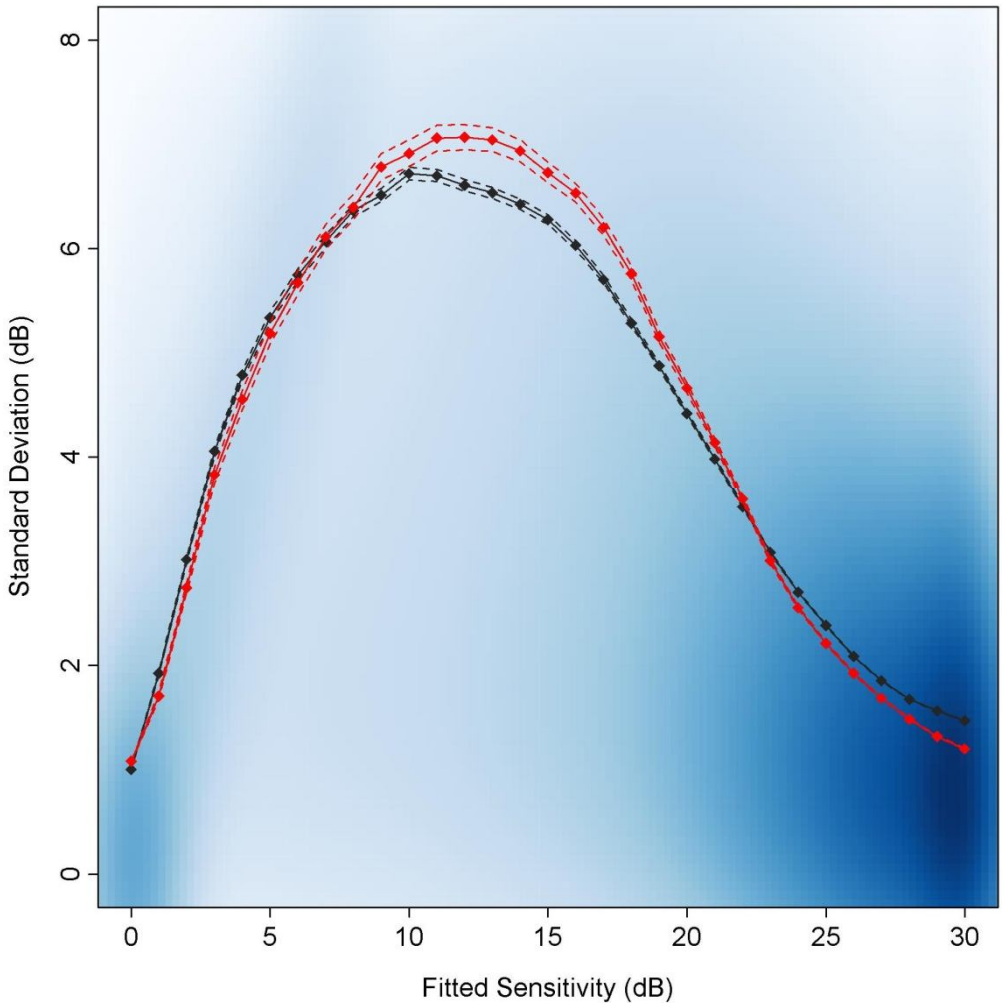

Figure 4.2 - Variability across sensitivities for SITA Standard (grey) and Fast (red). Solid lines represent standard deviation (SD) of residuals at each sensitivity and dashed lines 95% confidence intervals (CI) around those estimates of SD. The solid lines therefore represent the variability in SITA Standard and Fast, while the CI simply show these lines are not likely to be different due to chance. It is worth noting the widths of these intervals are slightly narrower for SITA Standard because there is more data and thus greater certainty about the SD estimates. The blue background is a smoothed colour density representation of the scatterplot for all residuals (approximately 4.5 million).

4.2.2. Simulated time to detect progression using SITA Standard and Fast testing algorithms

The results of 10,000 simulations of pointwise VF progression demonstrated that there was no meaningful difference in detection times using the two algorithms

(**Table 4.2**). For example, progression set at a speed of -1 dB per year at initially healthy levels of VF sensitivity is detected equally well by both the Fast (median: 4 years) and Standard algorithms (median: 4 years). **Figure 4.3** shows how this difference in precision between the SITA Standard and Fast algorithms might look in real life; in this figure the eye was simulated to progress at a constant rate across all points in the VF – variability was then empirically added according to the distributions of residuals for each true sensitivity level. There is little difference to see in the appearance of the greyscales of the VFs. This figure therefore illustrates the notion that any moderate additional variability associated with SITA Fast is barely noticeable.

Table 4.2 - Time to detect progression at p < 0.01 for SITA Standard and Fast using simulations of 100,000 progressing thresholds

Baseline Sensitivity (dB)	**Progression Rate (dB/year)**	**Median time to detect progression in years (Interquartile Range)**	
		SITA Standard	**SITA Fast**
10dB	**-0.5**	***14*** (3 to 18.5)	***14*** (3 to 18.5)
	-1	***9*** (3 to 12)	***9.5*** (3 to 12)
	-2	***6.5*** (3 to 9.5)	***6.5*** (3 to 10)
20dB	**-0.5**	***12.5*** (3.5 to 18)	***13*** (3.5 to 18.5)
	-1	***8.5*** (3 to 11.5)	***8.5*** (3 to 12)
	-2	***5.5*** (2.5 to 7.5)	***5.5*** (2.5 to 8)
30dB	**-0.5**	***6*** (3 to 8.5)	***5.5*** (2.5 to 7.5)
	-1	***4*** (2 to 5.5)	***4*** (2 to 5)
	-2	***3*** (1.5 to 4)	***2.5*** (1.5 to 3.5)

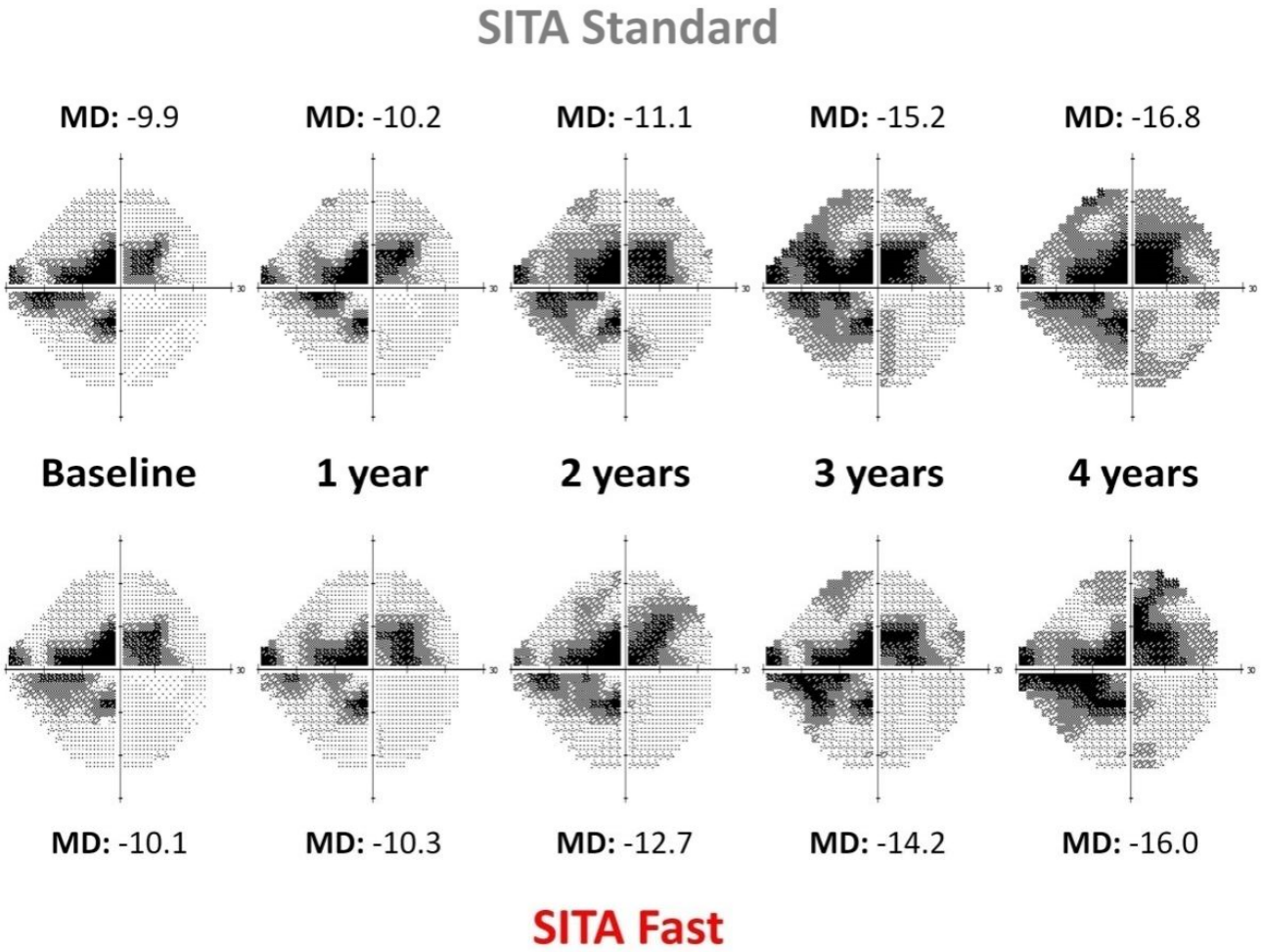

Figure 4.3 - Simulated greyscales produced from a baseline real life visual field of the same eye, but the variability from SITA Standard and Fast using R statistical software (R Development Core Team 2014). Each test location was simulated to progress at 2dB per year with noise added from the distributions of residuals for each fitted sensitivity.

4.3 Discussion

Consistent with previous studies, the variability of measurements generated using both SITA Standard and SITA Fast algorithms was found to increase with decreasing sensitivity for both tests before declining approaching perimetric blindness (Artes et al. 2002, Russell et al. 2012a). One reason why variability does not continually increase with VF loss is floor effects resulting from no measurements being possible below 0dB; this limitation of SAP seems to have some impact from below 20dB (**Figure 4.1**).

SITA Fast has slightly worse precision overall when compared to SITA Standard, though this difference is negligible at test locations with little or no sensitivity loss; in fact SITA Fast seems to be marginally more precise at this stage (perhaps due to initial presentations being closer to true thresholds). The difference in precision between methods is more apparent when thresholds drop to levels closer to 12dB and largest between around 15dB to 10dB, which matches the findings of Artes et al (Artes et al. 2002). Since variability differences between the two tests are low at near-normal thresholds, the shorter test times associated with SITA Fast may be preferable in glaucoma suspects or patients with early VF loss. Furthermore, the simulation results seem to indicate that any modest sacrifice in precision from using SITA Fast makes little difference to the time to detect pointwise progression with linear regression (**Table 4.2**). It is important to note that research suggests that a 30-60% drop in variability is required in SAP in order to make a significant difference to the speed at which deterioration is diagnosed (Turpin & McKendrick 2011); this level of difference was certainly not apparent in this study. On the other hand, the reduced test time of SITA Fast relative to SITA Standard (roughly two to three minutes according to the literature [Nordmann et al. 1998, Wild et al. 1999, Pierre-Filho et al. 2006]) may not be sufficient to make using this test worthwhile in the clinic.

Analysis of average times to detect different rates of VF deterioration, using a published model for simulating VF progression (Russell et al. 2013), was very informative (**Table 4.2**). When the level of pointwise VF damage drops below

20dB, the precision of the two algorithms are poor, perhaps too poor for detecting significant pointwise VF progression in a reasonable time period. For example, at that level of VF damage the model suggests it would take more than 8 years to detect a rate of -1dB per year with an examination every six months, whether using SITA Standard or Fast. This limitation about SAP has been well reported (Gardiner & Crabb 2002b, Chauhan et al. 2008b, Crabb & Garway-Heath 2012) and recently published data indicates that recording at these sensitivities, regardless of testing algorithm, may simply be unreliable; an apparent change in sensitivity within this range may not be informative of disease progression at all and should therefore be eliminated or counted as 0dB (Gardiner et al. 2014). If this is the case then differences in precision around this part of the measurement scale would be meaningless and it would be reasonable to conclude that SITA Fast could be as effective as SITA standard for follow-up.

4.3.1 Study Strengths and Limitations

This study did not consider the accuracy of the two testing algorithms. Evidence from computer simulation and real patient data suggest that both SITA strategies (although SITA Fast more so) tend to produce threshold values that are systematically higher than the original HFA Full-Threshold algorithm (Nordmann et al. 1998, Bengtsson & Heijl 1998b, Wild et al. 1999, Artes et al. 2002). Furthermore, given the reduction in number of tests required to calculate a threshold that, in each case, there is possibly a lower number of thresholds that SITA Fast can take. However, SITA Standard and Fast may not necessarily be less accurate than Full-Threshold testing, as this assumes that Full-Threshold is a gold-standard. Herein lies the issue: there is no gold-standard to evaluate the accuracy of these test algorithms; repeated testing of points to generate frequency of seeing curves (Gardiner et al. 2014) could be considered as one potential approach to calculate thresholds more accurately, but this could still be subject to non-constant variability across sensitivities, inter-patient variability and fatigue. Most importantly, VF total deviation measurements are age-corrected, so should not be affected by bias and given the similar scale of the rival methods, the relative accuracy of the tests should not affect utility in monitoring progression. In fact,

smaller variability in the SITA normative database has been shown to result in these algorithms actually being more sensitive to change than Full Threshold testing (Bengtsson & Heijl 1999). It is sometimes thought that, because both methods utilise surrounding thresholds to help derive sensitivity values and that SITA Fast tolerates a lower threshold of measurement certainty to reduce the number of stimulus presentations (Bengtsson & Heijl 1998b), VF damage has a higher likelihood of being missed or wrongly recorded in SITA Fast compared with Standard. This phenomenon was not tested in this study and can only be investigated in a prospective study.

One strength of the study was the fact that the precision estimates are based on millions of points from real clinical data rather than data from a small number of people taking part in a more controlled experiment. The sheer size of this dataset means that, whilst there are bound to be anomalies and outliers, the results are overall probably a good reflection of the precision of perimetric testing found in day-to-day clinical practice. A drawback with this approach is the lack of definitive clinical knowledge about the individual patients included in the study. For instance, patients may have had co-morbidities that affect the VF, like cataract, that could well affect estimates. In addition, even the type of diagnosis of glaucoma was not available – it is only known that patients were in glaucoma clinics. No exclusion criteria based upon reliability criteria (FPs, FNs and FLs) were applied to the datasets because much of these data was unavailable for this study. This is a limitation from the perspective that VFs with values for these indices above an arbitrary determined criterion are often excluded in clinical practice. However, it could be argued that these measurements are integral to the threshold estimates themselves. Further, one could also take the view that the reliability indices are possibly the most unreliable measures on the print-out (Shao et al. 2011) and this study makes the simplifying assumption that 'unreliable' VFs would have existed in equal measure in the Fast and Standard VFs, and this would not impact these findings. It is also important to acknowledge that these estimates of precision only represent an average rather than a measure that can be applied to different individuals; patients perform differently at perimetry and some produce

measurements with higher variability than others (Heijl et al. 1987, Russell et al. 2013). It was not possible to separate out variability caused by residual learning effects, technician experience, seasonal and time-of-day effects or any other sources of variability that may all have an effect on any given test. Without more detailed information on patient series there was no attempt to remove measured sensitivity outliers, as it was important not to introduce bias into the study with overly selective criteria with insufficient evidence. Furthermore, the distribution of cases is likely dominated by those with smaller amounts of VF loss and tests that are within normal limits, as the higher density of points in the bottom right corner of Figure 2 suggests. In addition it is important to note that these results estimate precision based upon following patients longitudinally rather than through obtaining test-retest data, so are probably more relevant in the context of looking at follow-up than the diagnosis of a single VF.

There are other aspects of the study that demand greater scrutiny. For instance, as a group, patients tested with SITA Fast were, on average, older than those tested using SITA Standard. It may be speculated that this reflects the preconception that older people are thought to be worse at perimetry (Ball et al. 1990, Chauhan & House 1991, Birt et al. 1997), which could lead to these patients being offered the quicker test in the belief that it is more suited to them. Also, the baseline MDs for SITA Fast tended to be higher than those for SITA Standard. This again may be explained by selection bias because the clinic databases would include people being followed without VF defects (patients at risk of developing glaucoma and ocular hypertensives) who may have been more likely to be assigned to the 'quicker' test. One assumption made in the study was that the model-fitted VF sensitivities were akin to patients' "true" sensitivities, which is reliant on assuming a linear relationship between change in VF thresholds and time. Alternative trend analyses modelling sensitivity changes over time have been proposed (Pathak et al. 2013, Azarbod et al. 2012, Caprioli et al. 2011, Russell & Crabb 2011), yet pointwise linear regression, as utilised here, is commonly used and, in practice, likely to be as effective (if not more) at predicting future loss as these other methods (McNaught et al. 1995, Bryan et al. 2013). In any case, for this study the regression is simply

used to extract an estimate of precision via the residuals. It is further worth noting that pointwise VF variability was analysed without regard for location in the VF even though it is already known that variability tends to differ according to test location (Heijl et al. 1987). However, this is confounded by the fact that sensitivity, which is not independent of variability, also varies according to test location (Henson et al. 2000, Artes et al. 2002, Russell et al. 2012a). As of yet, there is no clear evidence to suggest that there is a clinically significant difference in variability in test locations that is independent of sensitivity. It is worth noting that both SITA algorithms utilise 'smoothing', that is adjusting measured sensitivities according to adjacent locations and stimulus response times, at the end of testing (Bengtsson et al. 1997a); the 'noise' added to simulated progression could not include or take into account this additional source of variability, so the simulations are subject to slight oversimplification. It would perhaps also be interesting to look at the comparative performance of the testing algorithms using other criteria for defining loss in further work through simulations such as Glaucoma Change Probability Analysis (Bengtsson et al. 1997b, Heijl et al. 2003).

4.3.2 Further thoughts and Conclusions

Finally, it is most important to reflect on patient preference for different types of perimetry; it is imperative to account for this when choosing a testing algorithm. For example, some patients may feel fatigue more acutely than others and benefit from the shorter testing time of SITA Fast, whilst, for other patients it may be valuable to make use of the additional precision that comes with SITA Standard, particularly if the disease has progressed beyond an early stage. In the experience of clinical colleagues, some patients complain of having to delineate a high number of barely perceptible stimuli with SITA Fast because the algorithm presents stimuli closer to thresholds in order to reduce test time. Therefore, it is wrong to assume that SITA Fast is an 'easier' test. In practice it seems important to select the algorithm most suited to the patient and then, most essential of all, use this consistently throughout their follow-up (Musch et al. 2005), as clinical recommendations already stipulate (Chauhan et al. 2008b, European Glaucoma Society 2008).

Overall, this study suggests that there is some reduction in measurement precision associated with using SITA Fast instead of SITA Standard. However, this difference is unlikely to have a significant clinical impact on the time to detect VF progression. This study does not allow for any conclusion to be drawn about comparative accuracy of the measurements from SITA Standard and SITA Fast. However, regardless of which testing algorithm is used, it is clear that the time it takes to detect progression is perhaps too long to ensure that patients at risk of progressing to blindness are given timely intervention. As a result, it may be important to assess whether it is possible to anticipate future rapid progressors from any baseline characteristics. Investigation of the use of these risk factors in the context of anticipating future rates of loss will therefore be the subject of the next chapter.

Chapter Five: Using risk factors for fast glaucomatous progression to identify groups at risk of blindness

The ability to discriminate between fast and slow progressors of glaucoma would be clinically useful in ensuring that the appropriate level of treatment is given (Caprioli 2008). Determining the rate of progression as quickly as possible is therefore of high importance in glaucoma care. There is even evidence to suggest that rate of loss itself may impact on patient QoL (Lisboa et al. 2013), so it is essential these rates are controlled effectively. However, in the previous chapter, it was established that regardless of the testing algorithm utilised, detecting glaucomatous progression with any certainty using VFs alone can take a long time, perhaps too long to react in time to preserve patient visual function. This is obviously an important issue in determining VF deterioration. As a result, finding ways of singling out patients that are more at risk of rapid progression than others at baseline could potentially be clinically useful. Patients at risk of severe rates of progression could have their treatment intensity increased without need for the extensive follow-up as would be required if following patients solely utilising VFs. This chapter will therefore review the literature to determine whether patients at risk of fast progression can be identified more quickly by using other measurements taken in clinical practice, looking specifically at the utility of a risk calculator developed for this purpose by De Moraes et al. (De Moraes et al. 2012). This chapter ends with a study that demonstrates the limitations of this risk calculator.

5.1 Risk factors for fast disease progression

Risk factors associated with developing glaucoma, have already been looked at in Chapter 1. Many of these same risk factors have been linked to rate of VF progression. This section will hence proceed to go through the various risk factors found or thought to be associated with rates of functional loss.

5.1.1 Intraocular pressure

There exists a sizeable body of evidence that has linked the magnitude of measured

IOP to progression rate, observed in both retrospective analysis of clinical data (Teng et al. 2010, Heijl et al. 2012a, Chauhan et al. 2014) and in controlled clinical study (The Advanced Glaucoma Intervention Study Investigators 2000, Heijl & Bengtsson et al. 2009, Medeiros et al. 2012, Chauhan et al. 2014). There are a couple of instances where the relationship between IOP and progression rate is more ambiguous however. For instance, the Collaborative Normal Tension Glaucoma Study (CNTGS) looking at NTG patients found no significant relationship between progression rate and untreated IOP (Drance et al. 2001). However, this may be due to the fact that this study exclusively comprised of NTG patients, so, by definition, no subjects had notably high pressures.

There is some debate as to what IOP measures are most important in determining VF progression. Most of the time, average IOP over follow-up is utilised (The Advanced Glaucoma Intervention Study Investigators 2000, De Moraes et al. 2011a, Heijl et al. 2012a, Medeiros et al. 2012), although the peak IOP measurement over follow-up (De Moraes et al. 2011a) in addition to measurement of IOP at baseline (Heijl et al. 2009, Teng et al. 2010) have also been mooted as potentially useful. IOP variability measures such as the range (Heijl et al. 2012a, Chauhan et al. 2014) however, do not necessarily give a sense of the magnitude of IOP, which is generally important in determining risk. Regardless of how it is expressed, IOP has been consistently shown to have a profound effect on VF loss and a major predictive factor for progressing at rates that could lead to blindness.

5.1.2 Baseline Visual Field loss

As demonstrated in Chapter 3 (Saunders et al. 2014), initial VF damage at diagnosis is also likely to be linked to fast rates of progression thereafter. This is supported by Lee and colleagues, who found that those categorised as rapid progressors had significantly lower MDs than slower progressors (Lee et al. 2014). Further, in a cohort of patients with DH, Prata et al. estimated that eyes with a baseline MD worse than -4dB were over 200% more likely to be faster progressors (rate of loss worse than -1.5dB/year) (Prata et al. 2010). However, it is unclear whether the

correlation between VF damage at baseline and progression rate is because faster progressors are inevitably diagnosed at a later stage of disease, that rate of loss may change over the course of disease or whether this is a result of higher measurement accuracy at later disease stages due to staircasing effects (Malik et al. 2006). Contradicting these findings, Heijl et al., looking at patient records of Scandinavian patients, actually found that patients with a starting MD of better than -10.03dB had worse progression than those with earlier loss (Heijl et al. 2012a). It is, however, likely that ceiling effects were a large factor in this result given that a sizeable number of patients in the study had MDs worse than -20dB. To further confuse matters, Forchheimer et al. found, once IOP had been corrected for, no significant difference in progression rates between groups stratified according to baseline damage; the authors speculated that more intensive treatment with more advanced disease may have been a contributing factor in these findings (Forchheimer et al. 2011). Regardless, the fact that presentation time is likely to suggest something about how quickly patients are progressing should be given consideration.

5.1.3 Patient characteristics

Many studies have found significant relationships between age and progression rates. One study modelling progression rates estimated that the rate of VF deterioration increased by -0.019dB/year for each year lived (Heijl et al. 2012). Further, in the control arm of the EMGT, patients older than 68 progressed significantly more quickly than younger patients (Heijl et al. 2009), whilst the CGS also found increasing age was associated with worse rates of MD loss (Chauhan et al. 2010). Older patients have also been noted to progress more quickly than younger patients in retrospective data analysis (Prata et al. 2010, Heijl et al. 2012, Chauhan et al. 2014). Faster progressors furthermore tend to be older than slower progressing patients (Lee et al. 2014). However, not all studies have found a significant link between age and rate of VF loss. Utilising information from nine different population surveys and estimating rates of loss from prevalence rates and baseline WEMD Broman et al. could not find consistent evidence of age affecting rates of loss in OAG patients (Broman et al. 2008), while Drance et al. also could not

find evidence of age affecting progression rate in their CNTGS cohort (Drance et al. 2001). Overall, evidence seems to suggest that age is important to consider when analysing risk of fast progression, although this is perhaps offset by the fact that fast progression may not be as large an issue if residual life expectancy is lower.

Ethnicity is another factor that has been linked to rate of progression. Drance and colleagues found that Asians in the CNGTS had slower rates of progression than Caucasian patients, whilst the few enrolled black patients also tended to progress quickly (Drance et al. 2001). Similarly, Lee et al. found a higher proportion of black patients than Caucasian patients were faster progressors (Lee et al. 2014). In fact, Lee et al. found that African Americans had a 54% greater likelihood of being rapid progressors than Caucasians, although this was only after accounting for glaucoma severity at baseline. Finally, Broman et al. estimated that African patients tended to progress more quickly than Hispanic and White patients (Broman et al. 2008). Broman et al.'s results further suggest that Chinese patients are actually the fastest progressing group, seemingly contradicting the findings of the CNGTS, although this is potentially a result of far fewer patients being surveyed from this demographic than the others.

The CNGTS also found evidence that women tend to progress more quickly than men (Drance et al. 2001), but this has not been found consistently by a lot of other researchers (Broman et al. 2008, Prata et al. 2010, Chauhan et al. 2010, Heijl et al. 2012).

5.1.4 Structural factors

There are various structural indicators that have been linked to subsequent fast progression also. For instance, longitudinal optic disc change has been shown to precede measurable VF deterioration (Chauhan et al. 2009) and Lee et al. linked larger vertical cup-disc ratios to faster rates of VF loss (Lee et al. 2014). In addition, patients with focal optic disc damage have been seen to be faster progressors to patients with more diffuse damage (Reis et al. 2012). Given that structural damage will inevitably impact on functional loss it is perhaps not surprising that optic nerve

damage coincides with faster functional deterioration. As a result, many studies have attempted to use structural measurements to improve estimates of VF loss (Zhu et al. 2010, Medeiros et al. 2011, Russell et al. 2012b, Zhu et al. 2014) although this has often challenging due to the inherent variability in both sets of measurements (Gardiner & Johnson 2012).

Another characteristic linked to faster progression rates is thin CCT. Medeiros et al. found a significant link between CCT and rate of progression with each additional 100µm being associated with an improvement in prognosis of 0.21dB/year MD (Medeiros et al. 2012).

Myopia has also been linked to increased rate of VF loss in diagnosed glaucoma patients (Perdicchi et al. 2007), although myopia was not even observed as a risk factor for disease in one study looking at NTG patients (Sohn et al. 2010) leaving its status more ambiguous. The presence of Beta-zone parapapillary atrophy (Beta-PPA) has also been suggested to have some influence on rate of VF loss (Teng et al. 2010). This refers to degeneration of retinal pigment epithelial cells adjacent to the optic disc nerve head in addition to thinning of the choroidal or retinal cells in this area (see **Figure 5.1**), which is occasionally a feature of myopic eyes, but occurs more often (and tends to be larger) in eyes with glaucomatous optic-nerve damage (Jonas et al. 1989). Furthermore, in some studies, the development of this condition in ocular hypertensives and glaucoma patients seems to mirror the development of disease progression (Kono et al. 1999, Tezel et al. 2000). Speed of progression was not only found to be faster in eyes with Beta-PPA by Teng et al., but more eyes with MDs progressing at a faster rate than -0.5dB/year, had this optic disc characteristic (Teng et al. 2010).

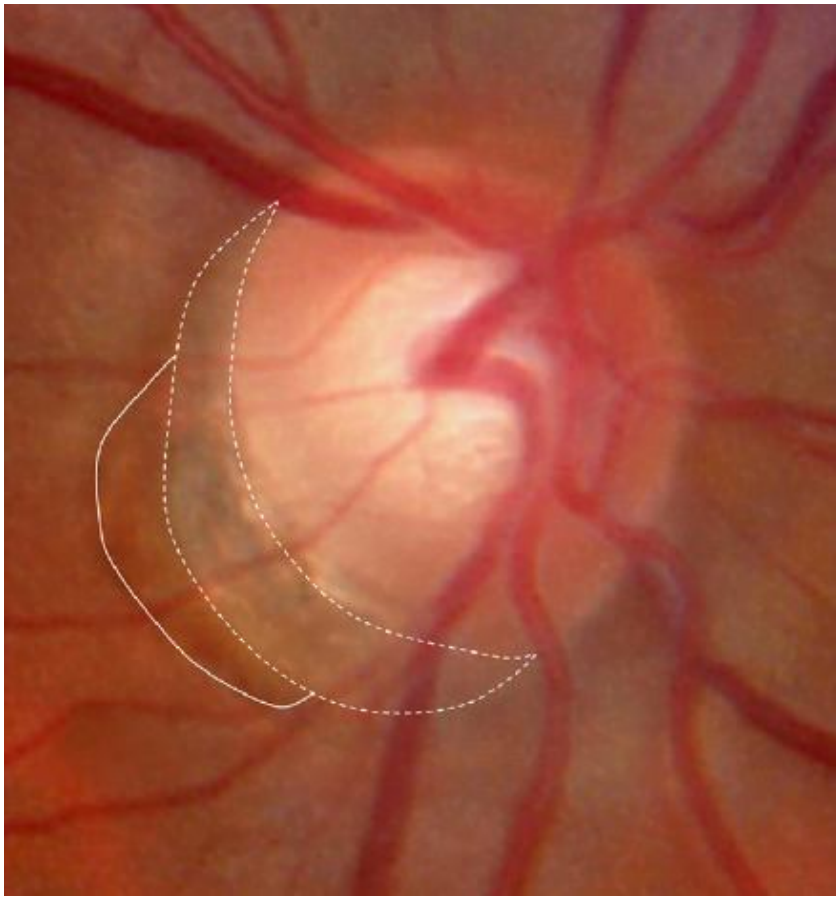

Figure 5.1 – An optic nerve – the solid and dashed lines show the locations of the alpha and Beta-zones respectively. Damage to retinal pigment epithelial cells in this area is defined as Beta-zone parapapillary atrophy. This image has been taken from Teng et al. 2010.

Other studies have found that specific features of glaucomatous loss may be linked to more rapid VF deterioration. For instance, Heijl et al. observed in the untreated arm of the EMGT that patients with pseudoexfoliation syndrome tended to have faster rates of loss than those that did not (Heijl et al. 2009). Pseudoexoliation glaucoma, however, was not found to be significantly faster when looking at data from clinical practice (Heijl et al. 2012). One big difference between the cohorts is the fact that patients in clinical practice are receiving treatment, so this could be one mechanism by which these contrasting results may have been observed. Alternatively, many pseudoexfoliative patients had higher levels of damage at baseline in the latter study, which may indicate that ceiling effects may have played some role in affecting the results.

Optic disc haemorrhages (DH) have also previously been linked to greater likelihood of VF deterioration (Drance et al. 1977, Siegner & Netland 1996, Leske et al. 2007, Kim 2014). These burst blood capillaries usually around the optic disc are often hard to find in an exam (no imaging devices can yet identify these automatically), but are a visible indicator that the optic disc is being damaged and that the disease is active (Radcliffe et al. 2014). Having long been recognised as a risk-factor for future VF damage or progression, there is evidence to suggest that it can have a profound effect on rate of progression also with one study showing that progression rates increased following DH (De Moraes et al. 2009a).

5.1.5 Other factors

In addition to the other potential risk factors for fast disease progression discussed above there are some circumstantial characteristics associated with faster disease progression worth mentioning. Surgery, be it laser or trabeculectomy, is one such factor. Heijl et al. found that more intensive treatment was associated with greater rates of progression, probably due to the fact that surgery is only performed when progression is not being controlled adequately. By definition, patients requiring surgery are therefore likely to be faster progressors (Heijl et al. 2012).

Drance et al. found some evidence that presence of migraine may be linked to rate of VF loss in patients with NTG, though it is worth noting that the majority of migrainous patients in their study were female. However, migraine was not found to be an important risk factor in a study investigating clinical data from patients with DH in clinical practice (Prata et al. 2010).

5.2 The De Moraes Risk Calculator

The ability to utilise these risk factors to anticipate which patients are at greatest risk of fast progression at diagnosis would be extremely useful clinically; potentially beneficial to both clinicians and patients in saving resources and time, whilst still

ensuring that patients receive the care they require and saving sight years. However, in spite of this, there have been few attempts to actually design a risk calculator for glaucoma patients.

The OHTS attempted to design a calculator for purposes of determining which patients with OHT are at greatest risk of progressing to glaucoma (Ocular Hypertension Treatment Study Group & European Glaucoma Prevention Study Group 2007). The use of the calculator developed has further been shown to be an improvement over clinical judgement in determining future development of glaucomatous disease (Mansberger & Cioffi 2006). However, this is of limited use for glaucoma patients given it was only designed for glaucoma suspects. Additionally, this calculator cannot give any information about whether the patient is at risk of progressing at a rate that may lead to blindness.

However, a group from New York recently developed two models using the retrospective New York Glaucoma Progression Study (NY-GAPS) to assess the probability of patient progression and the expected rate of this progression, validating their model utilising patient data from the Advanced Imaging in Glaucoma Study (AIGS) (De Moraes et al. 2012). The result of this is the first risk calculator developed for use in glaucoma.

5.2.1. Formulated model

There were two calculators developed: one for assessing risk of progression and another (of main interest in this section) focussing on predicting rate of MD loss per year, which will be hereafter referred to as the 'De Moraes model'. The variables obtained utilised the same variables derived from a previous study from the same group and included a number of continuous variables such as Age, CCT, peak IOP and mean IOP, some binary variables such as the presence of DH, Beta-PPA, exfoliation syndrome and whether the patient has had glaucoma surgery (De Moraes et al. 2011a). The De Moraes model developed using 587 eyes from 587 patients can be at the top of the following page:

$$Rate\ of\ VF\ change\ (dB/year) = -0.5343 + (Age) * -0.005227 + (CCT) * 0.002212 + (DH) * -0.16 + (Peak\ IOP) * -0.0259 + (mean\ IOP) * -0.008372 + (beta - PPA) * -0.04003 + (Exfoliation) * -0.07813 + (Glaucoma\ surgery) * 0.4521 + (Glaucoma\ surgery * mean\ IOP) * -0.04704$$

This model implies that faster rates of change are associated with higher age, lower CCT and higher peak and mean IOP, as one might expect. Considering the rough scales of the variables concerned CCT seems to be the most important variable with peak IOP and age significantly less important. Mean IOP is about a third as important as peak IOP. In addition, the presence of DH, Beta-PPA and exfoliation syndrome account for an extra -0.278 dB/year combined with DH the largest contributor. Glaucoma surgery, once the interaction term in the model is taken into account (considering the IOP will be over 10mm Hg), seems to suggest faster rates of loss, in agreement with previous studies (Heijl et al. 2012). As aforementioned, this is likely due to the fact that patients that require surgery are often the fastest progressors who most require the treatment in the first place.

5.2.2 Model evaluation methods

De Moraes et al. developed their model using the NY-GAPS database and evaluated it using VFs from AIGS. Their model was assessed using two methods: through comparing the fitted and observed rates of change in the validation dataset (AIGS), in addition to looking at the predicted and observed final MD values extrapolated from the calculated rates of progression (De Moraes et al. 2012). Observed rates of change were calculated from utilising linear regression of a whole series of points.

The performance of the model was then assessed using Bland-Altman plots (Bland & Altman 1986) comparing predicted against observed rates of MD change. These plots display the average and difference between two measurements, which is highly useful in assessing their agreement. In addition, the coefficient of determination (R^2), which measures the fit of the model, for a model predicting observed rates of loss was calculated.

5.2.3 The coefficient of determination – the R^2 and adjusted R^2 statistics

The well-known R^2 statistic, or the (multiple) coefficient of determination, pertains to the proportion of variance in the response variable explained by a fitted model relative to simply taking the mean of the response. In other words, it describes how well the model fits the data. An R^2 close to 1 implies an almost perfect relationship between the model and data, whereas an R^2 close to 0 implies that just fitting the mean is equivalent to the model fitted. Often, more than one variable is available to explain an outcome (multivariate model); in this situation, as variables are added, the R^2 will increase even if the variable is not important. The adjusted R^2 attempts to correct for this by penalising for increasing the number of variables, so is often preferred in comparing models. This statistic is often used in comparing models, but, unfortunately, there are no set criteria as to what universally represents a "good" adjusted R^2 value, so the only way of assessing such a statistic is via comparison with another predictive model.

5.2.4 Model evaluation results

The model comparing observed and predicted global rates of progression yielded an R^2 statistic of 0.14 and an adjusted R^2 of 0.13. The Bland Altman plot in **Figure 5.2** suggests that, on average, rate of loss estimates in the model tend to err towards being slightly pessimistic; predicted rates of loss average at 0.13dB/year worse than actual rates. The model tends to become more optimistic as observed progression rates worsen. The authors suggested, with these findings, that their risk calculator had "moderate accuracy". However, this assertion is, at best, baseless, and likely extremely misleading as the findings of the next section will suggest.

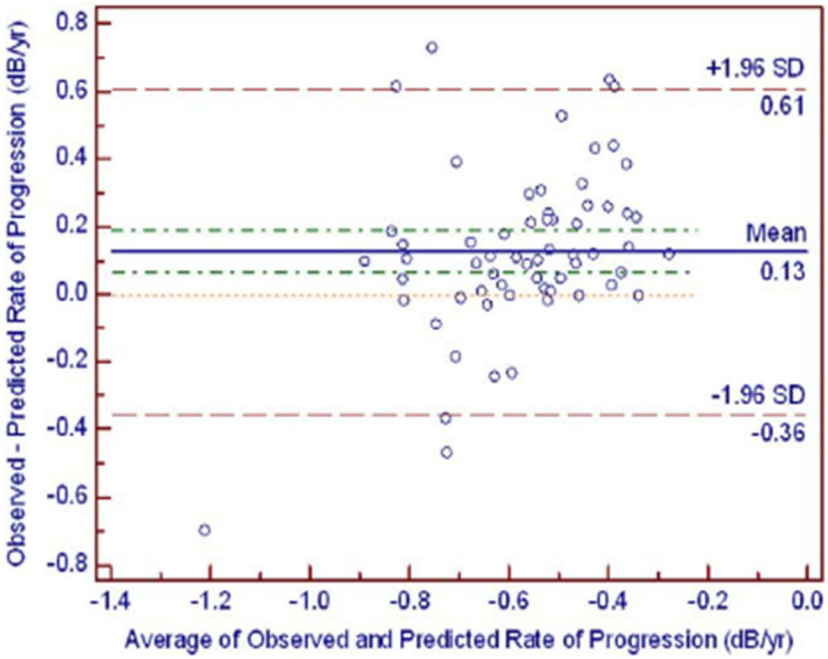

Figure 5.2 – This Bland-Altman plot displays the differences between observed and predicted rates of loss in the validation dataset taken from De Moraes et al. 2012

5.3 Evaluation of the model utility

It is difficult to calculate the utility of this calculator without comparison with a competitor model. For instance, it is uncertain whether this model, containing a whole battery of variables, has any more utility than simply making a prediction based upon, for example, the first two VFs. Although, it is not possible to use the same data as De Moraes et al. and direct comparison of R^2s is ill-advised, the work of the next section suggests that the usefulness of this model is not as high as the authors suggest. This work was published as a short article (letter) by *Investigative Ophthalmology and Visual Science* (Saunders et al. 2012b). Richard Russell (RR) and David Crabb (DC) were joint co-authors of this work and directed the project. RR further contributed to the data analysis, but I completed everything else described in the remainder of this chapter. All data utilised was collected from Moorfields Eye Hospital.

5.3.1 Methods

In order to illustrate that a model with an adjusted R2 of 0.13 is inappropriate for use as a predictive tool, an alternative, purely illustrative rate calculator using only the patient age (which has been shown to be related to rate of VF loss [Heijl et al. 2003]) and the patient's first two VFs was designed. This small number of variables contrasts to the De Moraes calculator, which used nine different variables.

To construct this model 68,099 anonymised VFs collected from 8,252 anonymised patients visiting the Glaucoma service at Moorfields Eye Hospital between 1997 and 2009 using the HFA (24-2 test pattern, Goldmann size III stimulus and SITA Standard testing algorithm) were utilised. Data were examined in accordance with the Declaration of Helsinki. Patients with fewer than 4 VFs (per eye) once the first VF was removed to account for learning effects were excluded. Furthermore, the interval between the first and second VFs was restricted to the range 3 to 13 months and this interval was not allowed to exceed 40% of the overall follow-up time. VF tests with FP or FN rates above 30% or FLs greater than 20% were discarded. Where both eyes of a patient were eligible, an eye was selected at random, leaving 875 patients (875 eyes) for investigation. The demographics of the study were comparable to the reference dataset in the De Moraes study, but their validation dataset contained patients with lower magnitude and variability of damage, as can be seen in **Table 5.1**.

Table 5.1 - A comparison of the demographics of datasets utilised in the De Moraes study (De Moraes et al. 2012) and the sample used for this project

	New York Glaucoma Progression Study	Advanced Imaging for Glaucoma Study	Moorfields Glaucoma service data
Number of Patients	587	62	875
Age at baseline (yrs)	64.9 ± 13.0	67.4 ± 8.3	62.7 ± 13.0
Baseline mean deviation (MD) (dB)	-7.1 ± 5.1	-3.7 ± 4.4	-7.0 ± 5.3
Follow-up time (yrs)	6.4 ± 1.7	4.0 ± 0.9	5.8 ± 1.7

The "true" rates of progression for each patient were calculated using ordinary least squares regression of the MD over time; the same method as that utilised in the De Moraes paper (De Moraes et al. 2012). In this model the difference between MDs in the second and first VFs divided by the time interval separating them was included to estimate an "initial rate of loss". The "true" rates of progression were then regressed against the baseline age of patients and their VF status across two visits as a basic model from which adjusted R^2 values were generated. All models fitted were of the form:

$$Rate\ of\ VF\ change\ (dB) = \beta_0 * Patient\ age\ (years) + \beta_1 * Initial\ Rate\ of\ VF\ loss\ \left(\frac{dB}{year}\right)$$

In order to facilitate comparisons, the study sample was split into a reference dataset and a validation dataset comprising exactly the same numbers of patients as included in the De Moraes study (i.e., 587 patients in the reference dataset and 62 patients in the validation dataset). To gain a distribution of values for the adjusted R^2 statistic, the 875 patients were randomly sampled without replacement 100,000 times into reference and validation datasets in order to attain 100,000 adjusted R^2s for each model (i.e., for the reference dataset, 587 patients were selected at random from the 875 patients in the complete dataset, and 62 patients

were sampled from the remaining 288 patients, and this was repeated 100,000 times). R^2 statistics were therefore obtained for both the reference datasets and models derived from the reference datasets applied to the validation datasets.

5.3.2 Results

The distribution of adjusted R^2s for the 100,000 reference models can be seen in **Figure 5.3**. The median adjusted R^2 is 0.10, whilst the reported R^2 for the De Moraes calculator lies at the 87th percentile (0.13). However, given that the reported R^2 could actually take any value between 0.125 and 0.135, the possibility of getting this statistic by chance in the Moorfields dataset could, in fact, be as high as 20%.

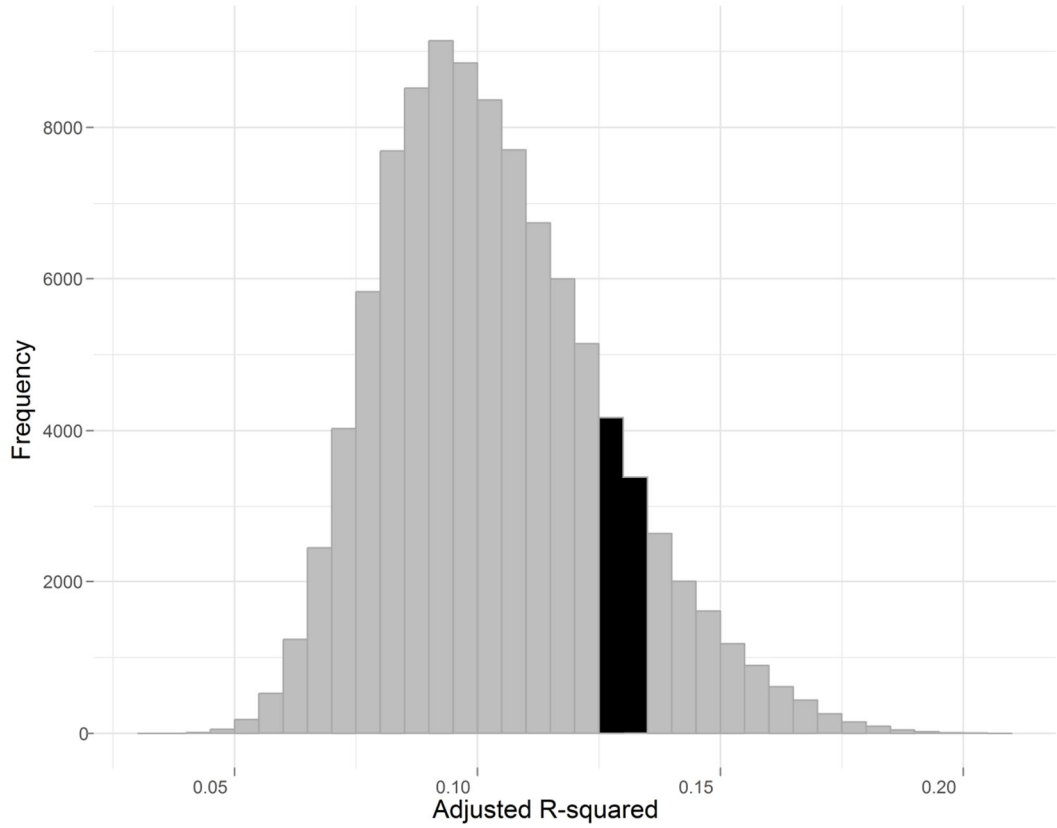

Figure 5.3 - A histogram showing the distribution of adjusted R^2 values from 100,000 simulated reference models. The black bars represent the potential range of R^2 values found by De Moraes et al. 2012.

Figure 5.4 shows the adjusted R^2 statistic yielded when the reference model is fitted to the validation dataset. The median adjusted statistic here is 0.08, but the

spread of this distribution should be noted; it was possible to simulate an adjusted R^2 statistic as high as 0.59 (due to the small sample size). The probability of gaining a better statistic than that of the De Moraes model was close to 35%.

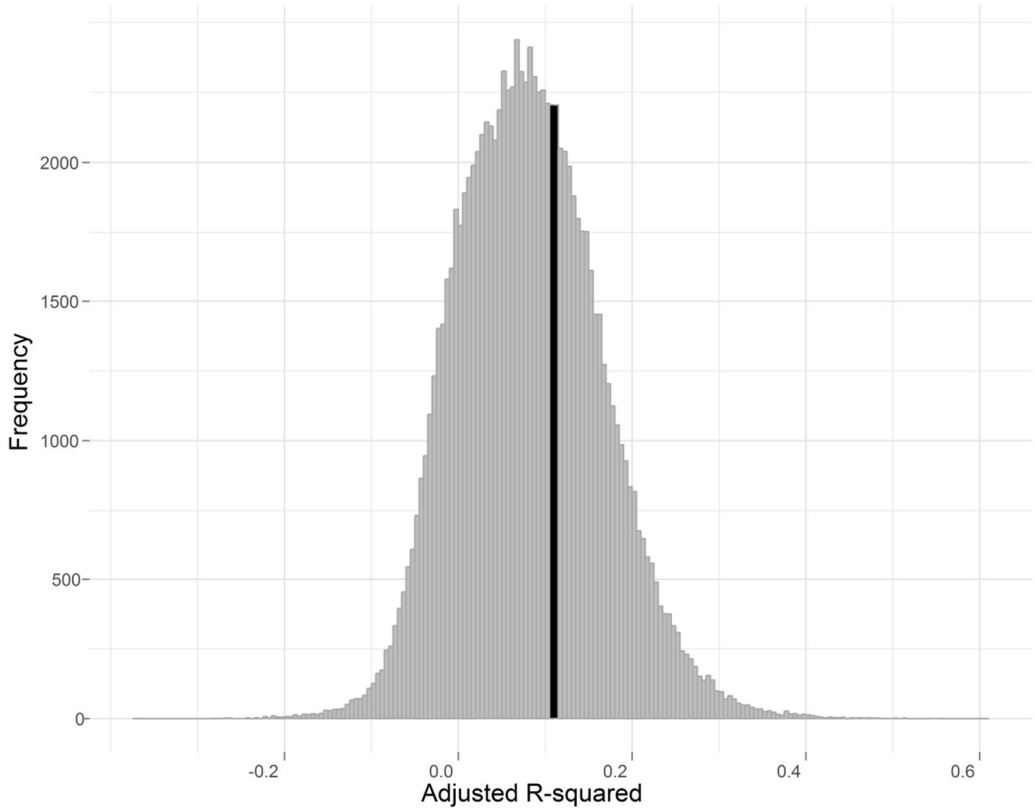

Figure 5.4 – A histogram showing the distribution of adjusted R^2 values from 100,000 simulated reference models fitted to simulated validation datasets. The black bars represent the potential range of R^2 values found by De Moraes et al. 2012.

One reference model was selected from the distribution in **Figure 5.3** with an R^2 similar in magnitude to that of the De Moraes rate calculator. The fit of this model can be seen in **Figure 5.5**, whilst **Figure 5.6** shows the effect of applying this model to a sample validation dataset (once again sampled to match the R^2 in De Moraes et al.'s paper); the 95% limits of agreement are shown by the dotted lines in **Figure 5.6**, and are more reflective of the likely range of differences between the estimated and actual rates of progression than the 95% CI for the average difference (indicated by the dashed lines) reported in the De Moraes et al. paper abstract. **Figure 5.5** and **Figure 5.6** clearly demonstrate the inadequacy of this

model, designed to mirror that of the De Moraes paper, for predicting rates of VF loss in spite of statistical significance.

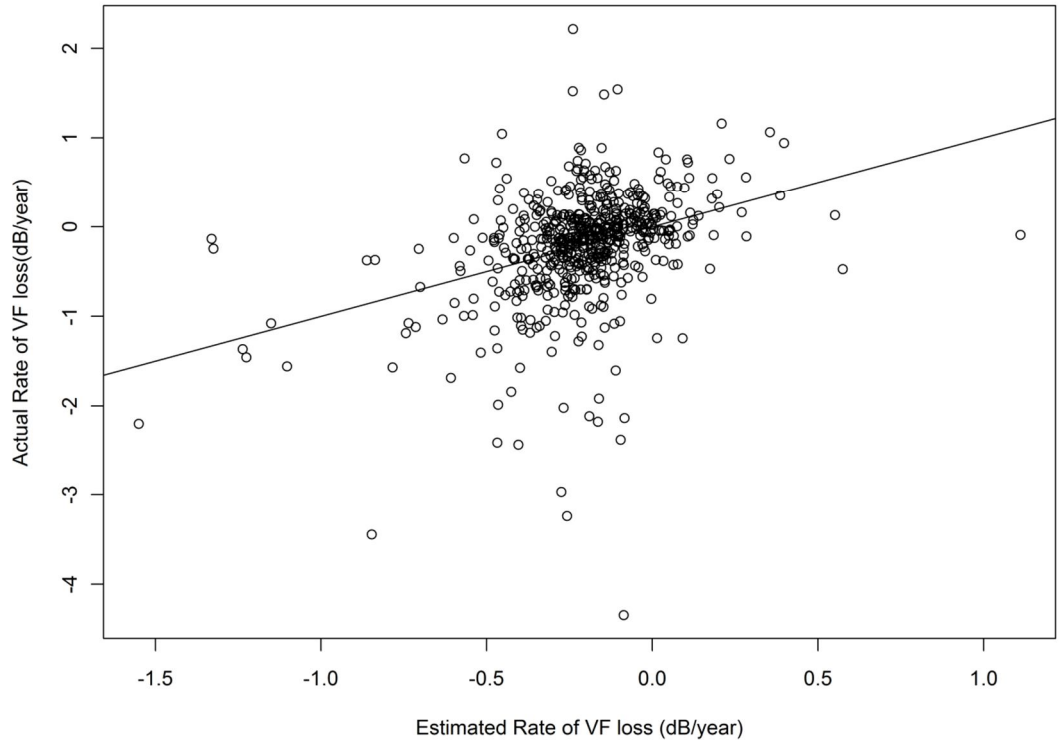

Figure 5.5 - A plot of the estimated progression rate of patients in the selected reference dataset against their "true" rate of progression. The solid line represents exact correspondence between the estimated and actual rates of progression (i.e. the line of unity). The R^2 is equal to 0.14 for this model.

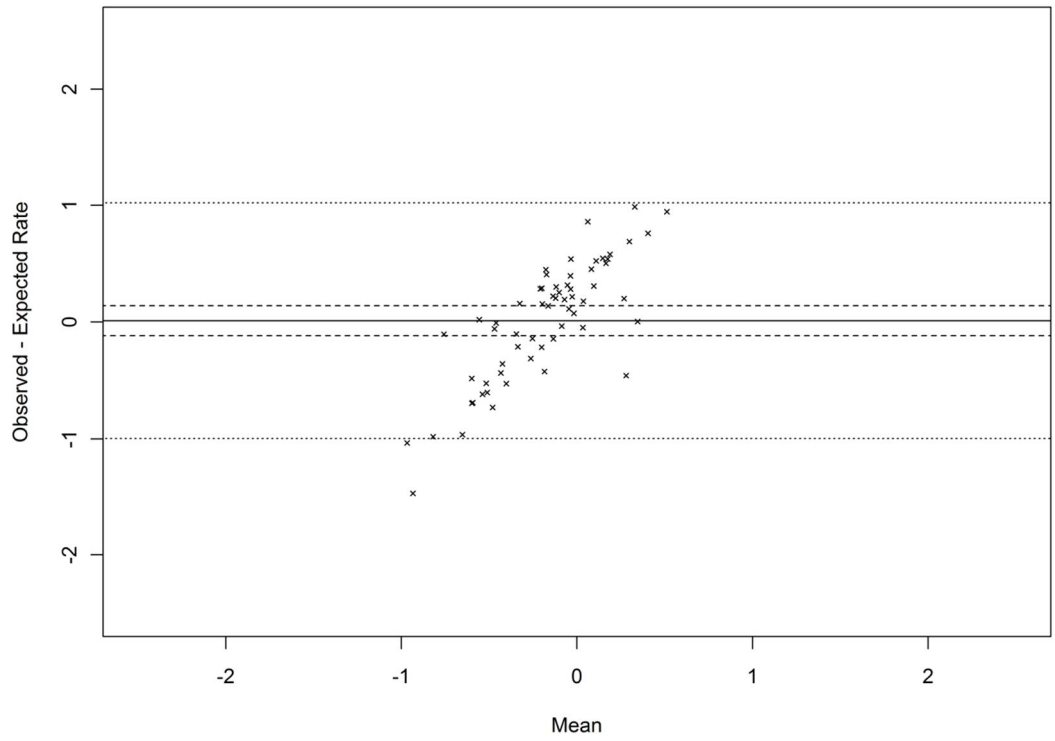

Figure 5.6 - A Bland-Altman plot comparing the progression rates of a selected validation dataset to the rates estimated in the model shown in figure 5.5 (R^2=0.12). The 95% Mean Confidence Interval (dashed lines) is 0.14 to -0.11dB per year. However, the more informative 95% limits of agreement (dotted lines) range from -1.0 to 1.02. As in De Moraes et al.'s model, the fit is worse at larger rates of progression.

5.3.3 Discussion

The "risk" model developed here would be poor in practice, and it is intuitively a bad idea to estimate a patient's rate of loss from their baseline damage and age alone. It does, nonetheless, give a sense of how poorly the De Moraes calculator predicts rate of loss. Though it is true that, strictly, the sets of R^2 values from this model are not directly comparable due to the fact that the data are different, the purpose of this exercise was to give some perspective to how the R^2 values presented by De Moraes et al. corresponded with a much more basic modelling method. It can be said that this very basic model performed almost as well with

modelling the progression rates of patients in the Moorfields dataset as the more complex De Moraes calculator did with data from the NY-GAPS and AIGS trials.

This study attempted to create a context for the De Moraes calculator's ability to estimate rate of VF loss, but it is important to note there are several limitations in the approach used in this study. For a start, the "validation" dataset utilised was not a different dataset from a different clinical centre and these patients had worse average MD than those used in validating the De Moraes calculator. Further, MDs of greater severities have been shown to be more variable than healthy ones (Russell et al. 2013), which suggests that rates of loss will be predicted less accurately in patients with more advanced VF damage. Thus, the characteristics of the De Moraes validation dataset may lead to more favourable results, as in the simpler model shown (see **Figure 5.6**), their calculator more accurately models patients with slower rates of VF loss.

There is another important message in this short study. R^2 statistics are often well understood and correctly interpreted, but can also be misleading, as the precision of the statistic is dependent on sample size (Wishart et al. 1931, Olkin & Finn 1995) and the coefficient is commonly presented without CIs or limits of tolerance. Without a sense of comparison the adjusted R^2 statistic is limited in its usefulness and it is apparent that there is, as yet, no reference standard for testing rate calculators such as this. The competitor model constructed in this study is very sensitive to the sample that was chosen, which is apparent from both **Figure 5.3** and **Figure 5.4**; the interquartile range of adjusted R^2s for the model just through selecting different field series could be as much as 10%. It is worth noting that the similar data characteristics and properties of the De Moraes dataset make it likely that their model is likely subject to similar sampling error, which suggests its utility is highly uncertain. **Figure 5.5** and **Figure 5.6** suggest that rate calculators with small R^2 values are inadequate for accurately predicting rates of loss especially in patients with fast progression, who are the most at risk of visual impairment.

Finally, it is important to emphasise that this illustrative rate calculator is not a serious attempt at introducing an alternative modelling strategy and should not be used to estimate rate of VF loss, yet it still seemed to provide similar predictive accuracy to the De Moraes calculator. The De Moraes calculator employed measurements that can be taken at the first visit, although the inclusion of glaucoma surgery as a baseline predictive variable is controversial in this context. In fact, given the substantial model coefficient associated with surgery, it would be interesting to see how their rate calculator would perform without this information. De Moraes et al. should be commended for their novel attempt at developing a statistical model for predicting progression in patients with treated glaucoma and especially for attempting to validate it using independent patient data. However, it would appear that the conclusion that must be drawn is that the limitations and low accuracy of their model make it completely unsuitable for clinical practice.

5.4 Conclusions

Although it would be incredibly helpful to be able to anticipate which patients are at greatest risk of fast progression, current risk calculators may not have the capability of sufficiently stratifying patients in this manner to be useful in clinical practice. Clearly the variables included in the De Moraes calculator are all associated with the rate of loss and the methods used to select them were correct, but this does not necessarily translate to development of a model with great utility. It is possible that some of the variables in the model could be linearly dependent (particularly the IOP peak and mean measurements) and that possibly a simpler model with less variables than the one proposed in the paper would have yielded similar results. Ultimately, until more is understood about how risk factors relate to progression and the mechanisms underlying disease, any risk calculator developed is inevitably going to suffer from issues of this nature.

As a result, it is still seems sensible to establish rate as quickly as possible. It has previously been suggested that testing six times within an initial interval of two years would particularly help establish this (Chauhan et al. 2008b). Accuracy of

measurements can be further optimised by grouping measurements at the beginning and end of the initial two year follow-up period (De Moraes et al. 2011b, Crabb & Garway-Heath 2012). However, the fact remains that this is not practiced clinically (Fung et al. 2013, Malik et al. 2013) and many clinicians judge that it is too large burden to place on time and resources (Malik et al. 2013). One potential solution, which would potentially help relieve this burden would be to use adaptive testing, adjusting intervals between tests and treatment according to change or variability in patient VFs (Jansonius 2006). Perhaps combining clinical information (such as IOP, CCT and presence of disc haemorrhage etc.) with VF measurements in a modelling context could further aid in the ability to assess the risk of patient progression (Medeiros et al. 2012). Overall, although baseline measurements are useful, for the moment, only time will tell whether a patient is one of the few that is in danger of progressing at a rate that could lead to visual impairment.

Chapter Six: Conclusions and Further Work

Though much research has been directed towards monitoring glaucomatous VF loss and, to an increasing extent, the impact of glaucoma on vision-related QoL (Glen et al. 2011), there are still many unanswered questions concerning what severity of disease impacts on QoL, how it can be best monitored and which patients are most at risk of progressing to visual impairment or blindness. The masses of VF data acquired in clinical practice are a useful resource in helping answer these questions and generating hypotheses for future studies. This thesis and the studies it contains has sought to utilise this UK clinical data for precisely this purpose. Below, I summarise the findings of the previous four chapters and the conclusions drawn from them before summarising the novel contributions of the thesis and describing further questions arising from this research.

6.1 Summary

The study described in Chapter 2 investigated what levels of VF severity are associated with loss of legal fitness to drive; an important landmark for patients in terms of independence (Fonda, Wallace et al. 2001). The IVF was used as a surrogate of a binocular test and legal VF requirements were used to categorise patients as individuals who would "pass" or "fail" the legal VF component of fitness to drive in the same manner set out by Crabb et al. (Crabb et al. 2004). Expected driving status was then compared against MDs recorded in clinical practice. The better eye MD was more useful at predicting whether patients would pass these criteria than the worse eye MD. Unsurprisingly, there was no utility from using the worse eye in predicting estimated legal fitness to drive. It was established that a criterion of worse than -14dB in the better eye was related to a probability of around 90% of failing the driving test. Clinical measurements in glaucoma have often been difficult to align with real life function and are therefore not relatable to patients. This research was novel because it was the first to attempt to devise a general MD threshold relating to legal fitness to drive to aid clinicians in evaluating VF measurements in the context of how it may impact on QoL.

The study in Chapter 3 utilised one of the landmarks to visual disability found in the previous chapter, in addition to clinical definitions of statutory blindness (US Social Security Administration 2011). Using these landmarks, this study sought to estimate how many individuals were estimated to be in danger of progressing to these damage levels within their predicted remaining lifetime. MDs in each eye at death were extrapolated from global rates of loss estimated from linear regression of MD in both eyes and ONS life expectancies (Office of National Statistics 2011). The findings were that a small, but significant minority of patients are progressing at a rate that would lead to statutory blindness (around 1 in 20) although a further 10% of patients progressed to a point where there was a 9 in 10 chance in becoming legally unfit to drive. Furthermore, this study found that around 70% of patients that were expected to progress to blindness actually had an MD worse than -6dB (or at least a "moderate defect" on the H-P-A scale [Hodapp et al. 1993]) at baseline. Only 1% progressed to this stage when diagnosed earlier suggesting that lateness of presentation is a big factor in future prognosis. There are plenty of studies that have looked at rates of loss for glaucoma patients (Heijl et al. 2012) and many that have estimated numbers of patients that have progressed to clinically diagnosed blindness (Hattenhauer et al. 1998, Oliver et al. 2002, Chen 2003). Additionally, utilising life expectancy in clinical decision making has also been suggested previously (Wesselink et al. 2011). However, this is the first study to practically consider patient progression rates in the context of their expected remaining lifetime to extrapolate future VF status in both eyes of the patient to estimate their future visual function.

The study in Chapter 4 sought to compare SITA Standard and SITA Fast testing algorithms to determine the difference in precision in measurements using the two methods and whether such differences could have an impact on the time to detect progression evaluated using simulations. Ultimately, a small difference in precision between the SITA Standard and Fast testing algorithms was observed, but this did not translate to any meaningful difference between the algorithms. As a result, it was concluded that it probably did not matter which algorithm was used. On one hand, it would perhaps better to use the faster test, yet, in real terms, this is only

likely to save 2 to 3 minutes per eye based on past studies (Nordmann et al. 1998, Wild et al. 1999, Pierre-Filho et al. 2006) and it is yet possible there is an unsubstantiated reduction in accuracy using SITA Fast. Overall, it is perhaps best to choose the easiest test for the patient, but it is hard to know which test this is; although SITA Fast is quicker, some clinical colleagues have suggested that it may be more difficult to do, although, as far as I am aware, there is not yet evidence to show this. Ultimately, it was apparent that both tests have precisions that mean that it could still take too long to diagnose progression in the context of patient vision loss. Although, relative precision of SITA Standard and Fast has been looked at once previously (Artes et al. 2002) this analysis was performed in comparison to Full Threshold testing, which itself is not a perfect gold standard. This study by contrast estimated each algorithm's precision individually in the context of its own measurements and is the first to attempt to estimate the clinical impact of any estimated differences in precision.

In the context of the fact that the previous chapter demonstrated the time to detect progression may be too long, Chapter 5 explored the utility of a current risk-calculator to estimate rates of loss based on baseline clinical characteristics proposed by De Moraes et al (De Moraes et al. 2012), through formulating a simple (and not clinically useful) model with similar statistical properties. The study demonstrated that the utility of models with as low an R^2 statistic as the De Moraes model is likely to be low in clinical practice. Due to the nature of their study and the fact that coefficients of determination for different data cannot be directly compared, it is not even certain whether their method has utility over simple models using just age and initial MDs. It is clear that there is still a lot of progress still to be made on attempting to establish which patients are at risk of fast progression at baseline.

6.2 Thesis contributions

Overall, there are various contributions to the field that have come out of this thesis:

- In the absence of any reference standard for clinical measurements, it has established evidence-based landmarks for visual impairment (Chapter 2).
- It has supported, using clinical data, estimates for proportions of patients classified as "fast progressors" for glaucoma (Chapter 3).
- It has looked at patient progression rates in the context of their predicted remaining lifetime (Chapter 3).
- It has established a modern estimate, albeit based on modelling, of how many patients with glaucoma are likely to go blind from their condition (Chapter 3).
- It has supported previous findings that baseline damage is important to future prognosis (Chapter 3).
- It has established precision estimates for sensitivities for SITA Standard and Fast using longitudinal data (Chapter 4).
- In the debate of whether to use SITA Standard or Fast, it is the first to utilise simulations to evaluate practical clinical impact on time to detect progression (Chapter 4).
- It has supported previous findings that it takes a long time to detect progression with any degree of certainty (Chapter 4).
- It has contextualised current risk calculators and established that it is still insufficient for anticipating risk of progression without VF data (Chapter 5).

6.3 Further work

What follows is a short description of some studies and hypotheses that could be future work arising from the results described in this thesis.

6.3.1 Topics from Chapter 2

In Chapter 2, the severity of patient VF loss, measured using BEMD, was related to legal fitness to drive in the UK. However, this does not necessarily relate to actual fitness to drive and there has been much criticism and demand for more research relating to this (Westlake 2000). Some patients that fall short of driving standards can feel as though they are safe drivers; in other words they do not perceive a loss in visual function, apart from in the application of VF standards. The evidence

relating driving safety to VF loss is often spurious and inconsistent and more evidence of the link between motor collisions in the context of mileage and VF damage is required. One potentially interesting avenue to explore would be to calculate the relationship between observed and expected numbers of motor vehicle collisions at different levels of VF loss. One could do this using a control population, but alternatively a model developed to incorporate various risk factors could be utilised (Maycock & Lockwood 1993). Initial analysis conducted using patient questionnaire data on accident frequency in the United Kingdom Glaucoma Treatment Study (UKGTS) has not revealed a significant relationship between IVF mean sensitivity once other mitigating factors including mileage have been accounted for.

It would perhaps also be useful to similarly relate clinical measurements to other important visual function landmarks such as difficulty reading and danger of falls, both of which could have a massive impact on a patient's overall QoL (Aspinall et al. 2008). As QoL is an under-investigated topic in glaucoma research (Glen et al. 2011) it could also be of interest to investigate the relationship between trial progression status (i.e. the point at which glaucomatous progression is flagged in a clinical trial) and reported QoL or even rates of loss and QoL given the findings of Lisboa et al. (Lisboa et al. 2013). In this respect, future targets in glaucoma treatment could be aligned to a measureable reduction in visual disability rather than potentially arbitrary changes in VF measurements.

A topic that has been the subject of much investigation is impact of VF loss on reading speed (Crossland et al. 2005, Ramulu et al. 2009a, Burton et al. 2012, Ramulu et al. 2012). Subject to a good deal of work so far has been a project that attempted to discover which areas of the VF were most correlated with impaired silent reading speed. This study utilised 92 participants (54 bilateral glaucoma patients from Moorfields Eye Hospital and 38 visually healthy age-related controls) with good corrected binocular acuity and attempted to determine whether silent reading speed was impacted by different levels of VF loss. Overall, no clear relationship between damage and reading speed could be found, either between controls and patients or relating to level of VF loss, mainly owing to the low

numbers of patients it was possible to recruit for the study and also due to a low range of VF loss. However, there was a slight suggestion that the inferior-left corner of the paracentral binocular VF, thought to be used in the return-sweep, may be more related than other parts. This work, undertaken with Robyn Burton and David Crabb has been submitted for publication in the *Japanese Journal of Ophthalmology* and, at time of writing, is currently under review.

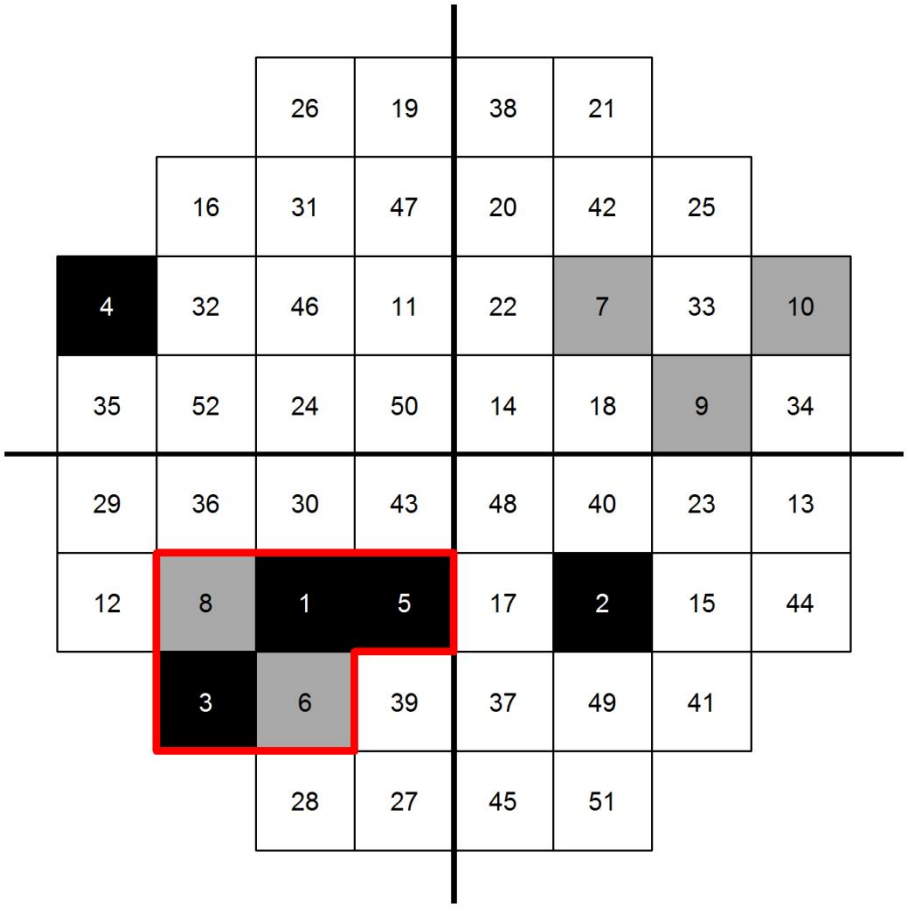

Figure 6.1 - A map displaying each test location in the binocular IVF ranked by R^2 statistics in the Reading study discussed in section 6.3.1. Black squares correspond with ranks 1 to 5, whilst grey squares represent ranks 6 to 10. The area bordered by red lines indicates a region that is likely to be of greater importance than other sections of the VF when considering reading speed for patients in this study. Figure taken from Burton et al. 2015.

6.3.2 Topics from Chapter 3

There were a number of interesting findings from Chapter 3, but there are plenty of questions arising from this research that merit extra investigation. For instance, although it was interesting to investigate the number of individuals in danger of progressing to a VF status where risk of visual impairment is high, there was no indication in this study of the number of sight-years lost due to the condition. It is obvious that being blind for one year is less devastating to a patient than being blind for ten, so it would be interesting to perform some further analysis to investigate how many years patients would be estimated to have to deal with impaired vision as a result of their glaucoma.

Further, though the study in Chapter 3 established that patients diagnosed with worse field loss at baseline are at greater risk of progressing to statutory blindness, it is not clear whether this was due to the fact that dangerously fast progressors are more likely to be diagnosed late or whether patients with advanced field loss at diagnosis progress similarly to those detected early, but are simply referred at too late a stage. The evidence for which scenario is the case is contrasting. Lee et al. reported that patients categorised as faster progressors had lower baseline MDs (Lee et al. 2014), yet Heijl et al., separating out their cohort using a median level of baseline damage, found that patients with an earlier stage of disease tended to progress more quickly than those with an MD over worse than -10dB at baseline (Heijl et al. 2012). However, due to the fact the study was not specifically tailored to investigate this, their results may have been affected by ceiling effects; an analysis stratifying starting levels of damage would be much more illuminating in answering this question. It would also perhaps be interesting to investigate whether rates of loss increase during the disease, although it would be difficult to do this without following a large cohort of patients for a very long period of time.

A recent study has suggested, reassuringly based on the results of this chapter, that patients, on average, are being diagnosed with less severe VF loss than 15 years ago, with NICE guidelines in particular thought to make a difference (Boodhna et al. 2014). One other opportunity this data affords is the ability to assess whether

rates of loss in glaucoma are changing over time. Potentially, if rates of loss are slower than in the past, this could change the outlook on patient prognosis forecasted.

Finally, it was notable in the study of Chapter 3 that there were many patients who progressed in one eye, but not the other. It would be interesting to investigate using large-scale clinical data how common this is in clinical practice.

6.3.3 Topics from Chapter 4

Chapter 4 seemingly established that using SITA Standard or Fast in patient follow-up made little difference, but one issue with this conclusion is that the study did not investigate the accuracy of the two tests. One possible consequence of the likely lower number of steps in the staircase algorithm as a consequence of requiring a lower level of certainty of the threshold before estimating it is that there is actually a lower range of measurements that a threshold can take using SITA Fast. An implication of this would be that potentially SITA Fast could have a tendency to fail to detect early, small VF scotomas. This has never been explicitly tested, although other studies seem to suggest the sensitivity of both algorithms of detecting early loss in practice is similar (Bengtsson & Heijl 1999, Pierre-Filho et al. 2006). Unfortunately, this could only really be assessed using a prospective study of test-retest data in a similar study design to that utilised by Artes et al. to measure the relative precision of the algorithms (Artes et al. 2002).

Another interesting question, given recommendations to give patients a test most suitable to them would be to have glaucoma patients of varying severity take the test and feedback which test is easier. This could be assessed utilising questionnaires or even through monitoring pupil activity directly as Henson et al. have done previously (Henson & Emuh 2010). This too would require a prospective cohort to investigate.

In the study of Chapter 4, simulations to evaluate time-to-detect progression, using noise from SITA Standard and Fast, were reliant on the assumption of VF progression being linear. It may therefore be interesting to utilise simulations of

other theoretical types of VF loss, including exponential regression and event-based analyses such as Glaucoma Change Probability Analysis.

Gardiner et al. suggest that there is so little utility in measuring thresholds below 19dB that these should be disregarded in analysis (Gardiner et al. 2014). In the study of Chapter 4, differences between SITA Standard and Fast were even less significant than they were in the current study. It would interesting to investigate whether findings from other studies would change through inserting a floor effect at this sensitivity threshold (i.e. rounding up thresholds below this cut-off to 19dB).

6.3.4 Topics from Chapter 5

Chapter 5 suggested that the only published risk calculator for predicting glaucomatous progression at baseline has an R^2 statistic that is likely not adequate in anticipating which patients are in danger of suffering long-term visual impairment. However, it would be useful to firmly establish this through direct comparison between the methods. The biggest flaw in the work by De Moraes et al. was the fact that they used an R^2 in isolation rather than through comparing their method with another current gold standard (which could have been as simple as simply utilising the initial two VFs). Ideally, it would be interesting to set up a study to find out the true efficacy of the De Moraes risk calculator against other standards for predicting future VF progression (for instance a simple model containing age and MD, or a short-term calculation of progression rate). What may be even more worth investigating would be developing a calculator incorporating patient age and baseline MD in calculating risk of progression to visual impairment. In the meantime, it is worth considering the use of a tool developed by Wesselink et al. (**Figure 6.2**), which attempts graphically to categorise the risk of an eye suffering severe impairment in a patient's lifetime from their MD, age and gender without knowing their rate of loss (Wesselink et al. 2011, Wesselink & Jansonius 2014). Estimates of risk have large uncertainty attributed to them, but this is sensible at present as it reflects the lack of information available at baseline. However, once there is sufficient follow-up to estimate rates of VF loss these can be

used to refine estimates of the risk of progressing to VF impairment (Wesselink et al. 2011).

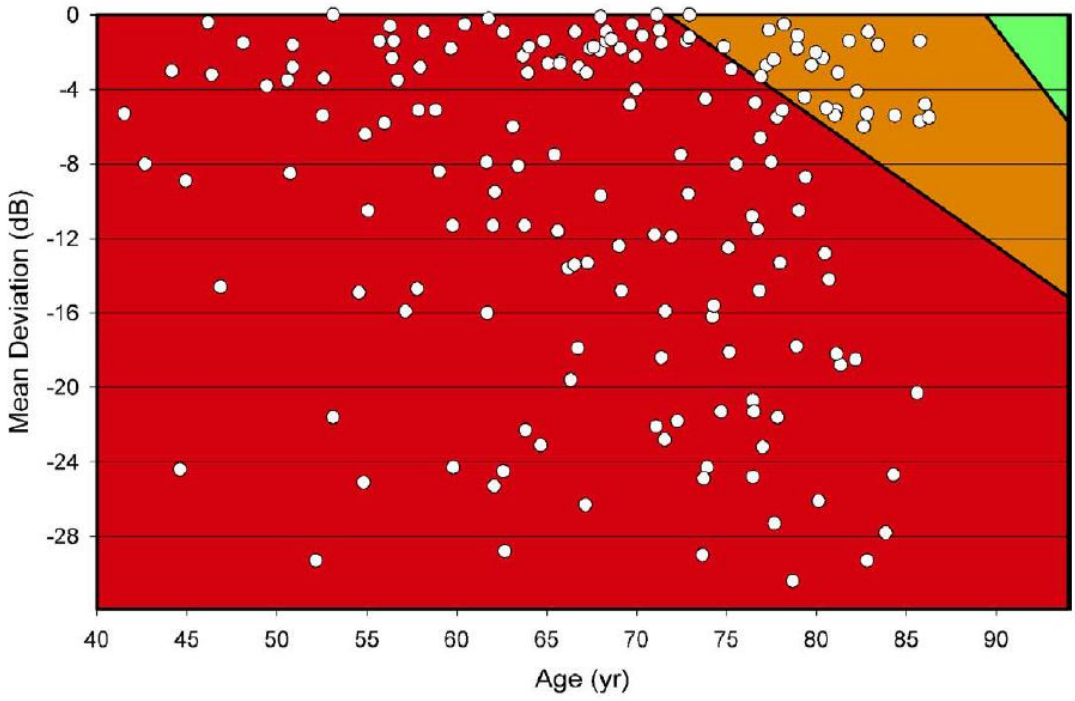

Figure 6.2 – A tool developed by Wesselink et al. 2011 to illustrate the risk of patients progressing to blindness from baseline for men (women have longer life-expectancies). The colours represent risk of becoming visually impaired (MD worse than -20dB) during a patient's predicted remaining lifetime. In the red zone, the probability of becoming visually impaired exceeds 2.5% even with treatment. In the amber zone, the probability of becoming visually impaired is below 2.5% with treatment, but greater than 2.5% without. Patients in the green zone have a lower than 2.5% chance of progressing to visual impairment even foregoing treatment. The white dots represent male patients from the Groningen Longitudinal Eye Study (Heeg et al. 2005). This figure was taken from Wesselink & Jansonius 2014.

6.3.5 Other questions

Using large amounts of retrospective data from clinical practice can potentially lend itself to a variety of other potential studies. For instance, it would be extremely useful from the perspective of many QoL studies to know the relative prevalence of superior and inferior VF loss. This would lend itself to being able to quantify how

many patients are likely to be in danger of suffering forms of visual impairment associated with losing vision in specific areas of the VF. For example, though Black et al. have established patients with inferior defects are at greater risk of falls (Black et al. 2011), it is unclear how many patients have predominantly inferior VF loss in glaucoma. In addition, it has been suggested that gender may be related to lateness of presentation for other diseases. So far in glaucoma, whether this is the case or not is ambiguous (Fraser et al. 1999, Fraser et al. 2001), so it would be interesting to use large-scale data to investigate this question.

Visual fields are an unusual measurement with many sources of variability, which means that there are plenty of pitfalls in their use in evaluating VF deterioration. However, massive amounts of clinical data can be used to learn more about the progression of glaucoma. Hopefully, in the near future, new methods could yet allow for earlier and more accurate glaucoma diagnosis and monitoring. For example, novel research aims to determine how structural defects can be better linked to functional ones to aid diagnoses with promising results thus far (Boland et al. 2008, Zhu et al. 2010). Incorporating Bayesian statistics into monitoring VF progression, using prior knowledge of patients with glaucoma or structural information, is another interesting research topic (Medeiros et al. 2011, Russell et al. 2012b), although this has only made a small difference so far. Another method has utilised the patient's own variability to define the likelihood of progression more clearly (O'Leary et al. 2012). Furthermore, various new modelling methods that overcome the difficulties of the assumptions of commonly-used linear regression methods being violated have been suggested (Caprioli et al. 2011, Russell & Crabb 2011, Pathak et al. 2013) However, one of the most promising methods of all thus far, called Analysis with Non-Stationary Weibull Error Regression with spatial enhancement (*ANSWERS*), has used patient test-retest data to model variability at different fitted sensitivities to inform modelling methods more accurately, increasing the speed of diagnosis and categorising eyes more effectively (Zhu et al. 2014). This Bayesian approach also incorporates structural information into its modelling structure. In summary, there are many new methods being developed to better monitor and predict future glaucomatous loss

and this large-scale clinical data is an excellent training-ground for testing new methods.

Finally, although not used in any studies in this thesis, the use of large-scale VF data could also potentially be useful for monitoring and auditing service delivery for glaucoma across clinical centres. This is already practiced in other ocular disease and VF data lends itself well to this endeavour. Further analysis of data collected from clinical practice may help suggest ways of improving the efficiency with which glaucoma is detected and managed in primary and secondary care and prevent avoidable sight loss.

Bibliography

THE ADVANCED GLAUCOMA INTERVENTION STUDY INVESTIGATORS, 1994. The Advanced Glaucoma Intervention Study - 2. Visual Test Scoring and Reliability. *Ophthalmology,* **101**(8), pp. 1445-1455.

THE ADVANCED GLAUCOMA INTERVENTION STUDY INVESTIGATORS, 2000. The Advanced Glaucoma Intervention Study (AGIS): 7. The Relationship Between Control of Intraocular Pressure and Visual Field Deterioration. *American Journal of Ophthalmology,* **130**, pp. 429-440.

ALTANGEREL, U.M.D., SPAETH, G.L.M. and RHEE, D.J.M., 2003. Visual function, disability, and psychological impact of glaucoma. *Current Opinion in Ophthalmology,* **14**, pp. 100-105.

ALTANGEREL, U.M.D., SPAETH, G.L.M. and STEINMANN, W.C., 2006. Assessment of Function Relation to Vision (AFREV). *Ophthalmic Epidemiology,* **13**, pp. 67-80.

ALTMAN, D.G. and BLAND, J.M., 1994a. Diagnostic tests 1: sensitivity and specificity. *British Medical Journal,* **308**, pp. 1552.

ALTMAN, D.G. and BLAND, J.M., 1994b. Diagnostic tests 2: predictive values. *British Medical Journal,* **309**, pp. 102.

ALTMAN, D.G. and BLAND, J.M., 1994c. Diagnostic tests 3: receiver operating characteristic plots. *British Medical Journal,* **309**, pp. 188.

ANG, G.S. and EKE, T., 2007. Lifetime visual prognosis for patients with primary open-angle glaucoma. *Eye,* **21**, pp. 604-608.

ARORA, K.S., BOLAND, M.V., FRIEDMAN, D.S., JEFFERYS, J.L., WEST, S.K. and RAMULU, P.Y., 2013. The Relationship between Better-Eye and Integrated Visual Field Mean Deviation and Visual Disability. *Ophthalmology,* **120**, pp. 2476-2484.

ARTES, P.H., IWASE, A., OHNO, Y., KITAZAWA, Y. and CHAUHAN, B.C., 2002. Properties of Perimetric Threshold Estimates from Full Threshold, SITA Standard, and SITA Fast Strategies. *Investigative Ophthalmology & Visual Science,* **43**(8), pp. 2654-2659.

ARTES, P.H. and CHAUHAN, B.C., 2005. Longitudinal changes in the visual field and optic disc in glaucoma. *Progress in Retinal and Eye Research,* **24**, pp. 333-354.

ARTES, P.H., CHAUHAN, B.C., KELTNER, J.L., CELLO, K.E., JOHNSON, C.A., ANDERSON, D.R., GORDON, M.O. and KASS, M.A., 2010. Longitudinal and Cross-sectional Analyses of Visual Field Progression in Participants of the Ocular Hypertension Treatment Study. *Archives of Ophthalmology,* **128**(12), pp. 1528-1532.

ARTES, P.H., O'LEARY, N., HUTCHISON, D.M., HECKLER, L., SHARPE, G.P., NICOLELA, M.T. and CHAUHAN, B.C., 2011. Properties of the Statpac Visual Field Index. *Investigative Ophthalmology & Visual Science,* **52**(7), pp. 4030-4038.

ARTES, P.H., O'LEARY, N., NICOLELA, M.T., CHAUHAN, B.C. and CRABB, D.P., 2014. Visual Field Progression in Glaucoma : What is the Specificity of the Guided Progression Analysis? *Ophthalmology*, **In Press.**

ASAOKA, R., CRABB, D.P., YAMASHITA, T., RUSSELL, R.A., WANG, Y.X. and GARWAY-HEATH, D.F., 2011. Patients Have Two Eyes!: Binocular versus Better Eye Visual Field Indices. *Investigative Ophthalmology & Visual Science,* **52**(9), pp. 7007-7011.

ASPINALL, P., A., JOHNSON, Z.K., AZUARA-BLANCO, A., MONTARZINO, A., BRICE, R. and VICKERS, A., 2008. Evaluation of Quality of Life and Priorities of Patients with Glaucoma. *Investigative Ophthalmology & Visual Science,* **49**(5), pp. 1907-1915.

ASRANI, S., ZEIMER, R., WILENSKY, J., GIESER, D., VITALE, S. and LINDENMUTH, K., 2000. Large Diurnal Fluctuations in Intraocular Pressure Are an Independent Risk Factor in Patients With Glaucoma. *Journal of Glaucoma*, **9**, pp. 134-142

AZARBOD, P., MOCK, D., BITRIAN, E., AFIFI, A.A., YU, F., NOURI-MAHDAVI, K., COLEMAN, A.L. and CAPRIOLI, J., 2012. Validation of Point-Wise Exponential Regression to Measure the Decay Rates of Glaucomatous Visual Fields. *Investigative Ophthalmology & Visual Science,* **53**, pp. 5403-5409.

BALL, K., OWSLEY, C. and BEARD, B., 1990. Clinical Visual Perimetry Underestimates Peripheral Field Problems in Older Adults. *Clinical Vision Sciences,* **5**(2), pp. 113-125.

BENGTSSON, B., OLSSON, J., HEIJL, A. and ROOTZÉN, H., 1997a. A new generation of algorithms for computerized threshold perimetry, SITA. *Acta Ophthalmologica Scandinavica,* **75**, pp. 368-375.

BENGTSSON, B., LINDGREN, A., HEIJL, A., LINDGREN, G., ASMAN, P. and PATELLA, V.M., 1997b. Perimetric probability maps to separate change caused by glaucoma from that caused by cararact. *Acta Ophthalmologica Scandinavica,* **75**, pp. 184-188.

BENGTSSON, B. and HEIJL, A., 1998a. Evaluation of a new perimetric threshold strategy, SITA, in patients with manifest and suspect glaucoma. *Acta Ophthalmologica Scandinavica,* **76**, pp. 268-272.

BENGTSSON, B. and HEIJL, A., 1998b. SITA Fast, a new rapid perimetric threshold test. Description of methods and evaluation in patients with manifest and suspect glaucoma. *Acta Ophthalmologica Scandinavica,* **76**, pp. 431-437.

BENGTSSON, B. and HEIJL, A., 1999. Comparing significance and magnitude of glaucomatous visual field defects using the SITA and Full Threshold strategies. *Acta Ophthalmologica Scandinavica,* **77**, pp. 143-146.

BENGTSSON, B., 2000. Reliability of computerized perimetric threshold tests as assessed by reliability indices and threshold reproducibility in patients with suspect and manifest glaucoma. *Acta Ophthalmologica Scandinavica,* **78**, pp. 519-522.

BENGTSSON, B. and HEIJL, A., 2008. A Visual Field Index for Calculation of Glaucoma Rate of Progression. *American Journal of Ophthalmology,* **145**(2), pp. 343-353.

BENGTSSON, B., PATELLA, V.M. and HEIJL, A., 2009. Prediction of Glaucomatous Visual Field Loss by Extrapolation of Linear Trends. *Archives of Ophthalmology,* **127**(12), pp. 1610-1615.

BHARGAVA, J.S., PATEL, B., FOSS, A.J.E., AVERY, A.J. and KING, A.J., 2006. Views of Glaucoma Patients on Aspects of Their Treatment: An Assessment of Patient Preference by Conjoint Analysis. *Investigative Ophthalmology & Visual Science,* **47**, pp. 2885-2888.

BIRT, C.M., SHIN, D.H., SAMUDRALA, V., HUGHES, B.A., KIM, C. and LEE, D., 1997. Analysis of Reliability Indices from Humphrey Visual Field Tests in an Urban Glaucoma Population. *Ophthalmology,* **104**(7), pp. 1126-1130.

BLACK, A.A., WOOD, J.M. and LOVIE-KITCHIN, J.E., 2011. Inferior Field Loss Increases Rate of Falls in Older Adults with Glaucoma. *Optometry and Vision Science,* **88**(11), pp. 1275-1282.

BLAND, J.M. and ALTMAN, D.G., 1986. Statistical Methods for Detecting Agreement Between Two Methods of Clinical Measurement. *Lancet,* **327**(8476), pp. 307-310.

BOLAND, M.V., ZHANG, L., BROMAN, A.T., JAMPEL, H.D. and QUIGLEY, H.A., 2008. Comparison of Optic Nerve Head Topography and Visual Field in Eyes with Open-angle and Angle-closure Glaucoma. *Ophthalmology,* **115**, pp. 239-245.

BOODHNA, T., SAUNDERS, L.J., RUSSELL, R.A. and CRABB, D.P., 2014. 'Are we getting better at detecting glaucoma before the disease becomes advanced? An examination of trends in visual field severity at the time of diagnosis', paper presented at *The Association for Research in Vision and Ophthalmology Annual Meeting*, Orlando FL, 6th May 2014.

BOZZANI, F.M., ALAVI, Y., JOFRE-BONET, M. and KUPER, H., 2012. A comparison of the sensitivity of EQ-5D, SF-6D and TTO utility values to changes in vision and perceived visual function in patients with primary open-angle glaucoma. *BMC Ophthalmology,* **12**:43.

BRANDT, J., GORDON, M.O., GAO, F., BEISER, J.A., MILLER, J.P., KASS, M.A. and the OCULAR HYPERTENSION TREATMENT STUDY GROUP, 2012. Adjusting

Intraocular Pressure for Central Corneal Thickness Does Not Improve Prediction Models for Primary Open-Angle Glaucoma. *Ophthalmology*, **119**(3), pp. 437-442

BROADWAY, D.C. and DRANCE, S.M., 1998. Glaucoma and vasospasm. *British Journal of Ophthalmology,* **82**, pp. 862-870.

BROMAN, A.T., QUIGLEY, H.A., WEST, S.K., KATZ, J., MUNOZ, B., BANDEEN-ROCHE, K., TIELSCH, J.M., FRIEDMAN, D.S., CROWSTON, J.G., TAYLOR, H.R., VARMA, R., LESKE, M.C., BENGTSSON, B., HEIJL, A., HE, M. and FOSTER, P.J., 2008. Estimating the Rate of Progressive Visual Field Damage in Those with Open-Angle Glaucoma from Cross-Sectional Data. *Investigative Ophthalmology & Visual Science,* **49**, pp. 66-76.

BRUSINI, P. and FILACORDA, S., 2006. Enhanced Glaucoma Staging System (GSS 2) for Classifying Functional Damage in Glaucoma. *Journal of Glaucoma,* **15**, pp. 40-46.

BRYAN, S.R., VERMEER, K.A., EILERS, P.H.C., LEMIJ, H.G. and LESAFFRE, EMMANUEL M. E. H., 2013. Robust and Censored Modeling and Prediction of Progression in Glaucomatous Visual Fields. *Investigative Ophthalmology & Visual Science,* **54**, pp. 6694-6700.

BURG, A., 1971. Vision and Driving: A Report on Research. *Human Factors,* **13**(1), pp. 79-87.

BURTON, R., CRABB, D.P., SMITH, N.D., GLEN, F.C. and GARWAY-HEATH, D.F., 2012. Glaucoma and reading: exploring the effects of contrast lowering of text. *Optometry and Vision Science,* **89**(9), pp. 1282-1287.

BURTON, R., **SAUNDERS, L.J.** and CRABB, D.P., 2015. Areas of the visual field important during reading in Glaucoma, *Japanese Journal of Ophthalmology,* **Online First**.

CAPRIOLI, J., 2008. The Importance of Rates in Glaucoma. *American Journal of Ophthalmology,* **145**(2), pp. 191-192.

CAPRIOLI, J., MOCK, D., BITRIAN, E., AFIFI, A.A., YU, F., NOURI-MAHDAVI, K. and COLEMAN, A.L., 2011. A Method to Measure and Predict Rates of Regional Visual Field Decay in Glaucoma. *Investigative Ophthalmology & Visual Science,* **52**, pp. 4765-4773.

CARL ZEISS MEDITEC INC, 2003. *Humphrey Glaucoma Progression Analysis (GPA) Software - An advanced approach to monitoring disease progression*, Jena: Carl Zeiss Meditec AG.

CARL ZEISS MEDITEC INC, 2008. *HFA II-i with Guided Progression Analysis (GPA) - Sample Cases*. Jena: Carl Zeiss Meditec AG.

CHAN, E.W., CHIANG, P.P.C., WONG, T.Y., SAW, S.M., LOON, S.C., AUNG, T. and LAMOUREUX, E., 2013. Impact of Glaucoma Severity and Laterality on Vision-

Specific Functioning: The Singapore Malay Eye Study. *Investigative Ophthalmology & Visual Science,* **54**, pp. 1169-1175.

CHAUHAN, B.C. and HOUSE, P.H., 1991. Intratest Variability in Conventional and High-pass Resolution Perimetry. *Ophthalmology,* **98**(1), pp. 79-83.

CHAUHAN, B.C., MIKELBERG, F.S., BALAZSI, A.G., LEBLANC, R.P., LESK, M.R. and TROPE, G.E., 2008a. Canadian Glaucoma Study: 2. Risk Factors for the progression of open-angle glaucoma. *Archives of Ophthalmology,* **126**(8), pp. 1030-1036.

CHAUHAN, B.C., GARWAY-HEATH, D.F., GONI, F.J., ROSSETTI, L., BENGTSSON, B., VISWANATHAN, A.C. and HEIJL, A., 2008b. Practical recommendations for measuring rates of visual field change in glaucoma. *British Journal of Ophthalmology,* **92**, pp. 569-573.

CHAUHAN, B.C., NICOLELA, M.T. and ARTES, P.H., 2009. Incidence and Rates of Visual Field Progression after Longitudinally Measured Optic Disc Change in Glaucoma. *Ophthalmology,* **116**, pp. 2110-2118.

CHAUHAN, B.C., MIKELBERG, F.S., ARTES, P.H., BALAZSI, A.G., LEBLANC, R.P., LESK, M.R., NICOLELA, M.T. and TROPE, G.E., 2010. Canadian Glaucoma Study – 3. Impact of Risk Factors and Intraocular Pressure Reduction on the Rates of Visual Field Change. *Archives of Ophthalmology,* **128**(10), pp. 1249-1255

CHAUHAN, B.C., MALIK, R., SHUBA, L.M., RAFUSE, P.E., NICOLELA, M.T. and ARTES, P.H., 2014. Rates of Glaucomatous Visual Field Change in a Large Clinical Population. *Investigative Ophthalmology & Visual Science,* **In Press**.

CHEN, P.P., 2003. Blindness in patients with treated open-angle glaucoma. *Ophthalmology,* **110**(4), pp. 726-733.

CHISHOLM, C.M., RAUSCHER, F.G., CRABB, D.P., DAVIES, L.N., DUNNE, M.C., EDGAR, D.F., HARLOW, A.J., JAMES-GALTON, M., PETZOLD, A., PLANT, G.T., VISWANATHAN, A.C., UNDERWOOD, G.J. and BARBUR, J.L., 2008. Assessing visual fields for driving in patients with paracentral scotomata. *British Journal of Ophthalmology,* **92**, pp. 225-230.

COECKELBERGH, T.R.M., CORNELISSEN, F.W., BROUWER, W.H. and KOOIJMAN, A.C., 2002. The effect of visual field defects on eye movements and practical fitness to drive. *Vision Research,* **42**, pp. 669-677.

CRABB, D.P., VISWANATHAN, A.C., MCNAUGHT, A.I., POINOOSAUMY, D., FITZKE, F.W. and HITCHINGS, R.A., 1998. Simulating binocular visual field status in glaucoma. *British Journal of Ophthalmology,* **82**, pp. 1236-1241.

CRABB, D.P., FITZKE, F.W., HITCHINGS, R.A. and VISWANATHAN, A.C., 2004. A practical approach to measuring the visual field component of fitness to drive. *British Journal of Ophthalmology,* **88**, pp. 1191-1196.

CRABB, D.P. and VISWANATHAN, A.C., 2005. Integrated visual fields: a new approach to measuring the binocular field of view and visual disability. *Graefe's Archive for Clinical and Experimental Ophthalmology,* **54**, pp. 1169-1175.

CRABB, D.P., SMITH, N.D., RAUSCHER, F.G., CHISHOLM, C.M., BARBUR, J.L., EDGAR, D.F. and GARWAY-HEATH, D.F., 2010. Exploring Eye Movements in Patients with Glaucoma When Viewing a Driving Scene. *PLoS ONE*, **5**(3): e9710

CRABB, D.P. and GARWAY-HEATH, D.F., 2012. Intervals between visual field tests when monitoring the glaucomatous patient: *Wait and See*. *Investigative Ophthalmology & Visual Science,* **53**(6), pp. 2770-2776.

CRABB, D.P., SMITH, N.D., GLEN, F.C., BURTON, R. and GARWAY-HEATH, D.F., 2013. How Does Glaucoma Look? : Patient Perception of Visual Field Loss. *Ophthalmology,* **120**(6), pp. 1120-1126.

CROSSLAND, M.D., CULHAM, L.E. and RUBIN, G.S., 2005. Predicting Reading Fluency in Patients with Macular Disease. *Optometry and Vision Science,* **82**(1), pp. 11-17.

CUMMINGS, R.W. and RUBIN, G.S., 1993. Reading speed and saccadic eye movements with an artifical paracentral scotoma. *Investigative Ophthalmology & Visual Science,* **34**, pp. 1318.

DE MORAES, C.G.V., PRATA, T.S., LIEBMANN, C.A., TELLO, C., RITCH, R. and LIEBMANN, J.M., 2009a. Spatially Consistent, Localized Visual Field Loss before and after Disc Hemorrhage. *Investigative Ophthalmology & Visual Science,* **50**, pp. 4727-4733.

DE MORAES, C.G.V., PRATA, T.S., TELLO, C., RITCH, R. and LIEBMANN, J.M., 2009b. Glaucoma With Early Visual Field Loss Affecting Both Hemifields and the Risk of Disease Progression. *Archives of Ophthalmology,* **127**(9), pp. 1129-1134.

DE MORAES, C.G.V., JUTHANI, V.J., LIEBMANN, J.M., TENG, C.C., TELLO, C., SUSANNA JR., R. and RITCH, R., 2011a. Risk Factors for Visual Field Progression. *Archives of Ophthalmology,* **129**, pp. 562-568.

DE MORAES, C.G.V., RITCH, R., TELLO, C. and LIEBMANN, J.M., 2011b. Modified Visual Field Trend Analysis. *Journal of Glaucoma,* **20**(4), pp. 203-206.

DE MORAES, C.G.V., SEHI, M., GREENFIELD, D.S., CHUNG, Y.S., RITCH, R. and LIEBMANN, J.M., 2012. A Validated Risk Calculator to Assess Risk and Rate of Visual Field Progression in Treated Glaucoma Patients. *Investigative Ophthalmology & Visual Science,* **53**, pp. 2702-2707.

DE VOOGD, S., IKRAM, M.K., WOLFS, R.C.W., JANSONIUS, N.M., HOFMAN, A. and DE JONG, PAULUS T. V. M., 2005. Incidence of Open-Angle Glaucoma in a General Elderly Population - The Rotterdam Study. *Ophthalmology,* **112**, pp. 1487-1493.

DELONG, E.R., DELONG, D.M. and CLARKE-PEARSON, D.L., 1988. Comparing the Areas Under Two or More Correlated Receiver Operating Characteristic Curves: A Nonparametric Approach. *Biometrics,* **44(3)**, pp. 837-845.

DRANCE, S.M., FAIRCLOUGH, M., BUTLER, D.M. and KOTTLER, M.S., 1977. The importance of disc haemorrhage in the prognosis of open-angle glaucoma. *Archives of Ophthalmology,* **95(2)**, pp. 226-228.

DRANCE, S.M., ANDERSON, D.R. and SCHULZER, M., 2001. Risk Factors for Progression of Visual Field Abnormalities in Normal-tension Glaucoma. *American Journal of Ophthalmology,* **131**, pp. 699-708.

DRIVERS MEDICAL GROUP, 2013. For Medical Practitioners - At a glance Guide to the current Medical Standards of Fitness to Drive. Swansea, UK: Drivers and Vehicle Licensing Agency.

ELBOZAN CUMURCU, B., CUMURCU, T., CAM CELIKEL, F. and ETIKAN, I., 2010. The Relation Between Quality of Life and Visual Functioning in Patients with Glaucoma. *Turkiye Klinikleri Journal of Medical Sciences,* **30**(2), pp. 698-704.

ERNEST, P.J.G., SCHOUTEN, JAN S. A. G., BECKERS, H.J.M., HENDRIKSE, F., PRINS, M.H. and WEBERS, C.A.B., 2011. The evidence base to select a method for assessing glaucomatous visual field progression*. Acta Ophthalmologica Scandinavica,* **90**(2), pp. 101-108.

ESTERMAN, B., 1982. Functional Scoring of the Binocular Field. *Ophthalmology,* **89**(11), pp. 1226-1234.

EUROPEAN GLAUCOMA PREVENTION STUDY GROUP, 2007. Predictive Factors for Open-Angle Glaucoma among Patients with Ocular Hypertension in the European Glaucoma Prevention Study. *Ophthalmology,* **114**, pp. 3-9.

EUROPEAN GLAUCOMA SOCIETY, 2008. *European Glaucoma Society Guidelines (2008): Terminology and Guidelines for Glaucoma.* 3rd edition. Savona, Italy: Editrice DOGMA.

FANKHAUSER, F., SPAHR, J. and BEBIE, H., 1977. Some Aspects of the Automation of Perimetry. *Survey of Ophthalmology,* **22**(2), pp. 131-141.

FITZKE, F.W., HITCHINGS, R.A., POINOOSAUMY, D., MCNAUGHT, A.I. and CRABB, D.P., 1996. Analysis of visual field progression in glaucoma. *British Journal of Ophthalmology,* **80**, pp. 40-48.

FLAMMER, J., 1986. The concept of visual field indices. *Graefe's Archive for Clinical and Experimental Ophthalmology,* **224**(5), pp. 389-392.

FONDA, S.J., WALLACE, R.B. and HERZOG, A.R., 2001. Changes in Driving Patterns and Worsening Depressive Symptoms Among Older Adults. *Journal of Gerontology,* **56B**(6), pp. S343-S351.

FORCHHEIMER, I., DE MORAES, C.G.V., TENG, C.C., FOLGAR, F., TELLO, C. and LIEBMANN, J.M, 2011. Baseline mean deviation and rates of visual field change in treated glaucoma patients. *Eye*, **25**, pp. 626-632

FRASER, S., BUNCE, C. and WORMALD, R., 1999. Risk Factors for Late Presentation in Chronic Glaucoma. *Investigative Ophthalmology & Visual Science,* **40**, pp. 2251-2257.

FRASER, S., BUNCE, C., WORMALD, R. and BRUNNER, E., 2001. Deprivation and late presentation of glaucoma: case-control study. *British Medical Journal,* **322**, pp. 639-643.

FREEMAN, E.E., MUNOZ, B., TURANO, K.A. and WEST, S.K., 2006. Measures of Visual Function and Their Association with Driving Modification in Older Adults. *Investigative Ophthalmology & Visual Science,* **47**, pp. 514-520.

FREEMAN, E.E., MUNOZ, B., WEST, S.K., JAMPEL, H.D. and FRIEDMAN, D.S., 2008. Glaucoma and quality of life: the Salisbury Eye Evaluation. *Ophthalmology,* **115**, pp. 233-238.

FUNG, S.S.M., LEMER, C., RUSSELL, R.A., MALIK, R. and CRABB, D.P., 2013. Are practical recommendations practiced? A national multi-centre cross-sectional study on frequency of visual field testing in glaucoma. *British Journal of Ophthalmology,* **97**(7), pp. 843-847.

GABRIELE, M.L., WOLLSTEIN, G., ISHIKAWA, H., KAGEMANN, L., XU, J., FOLIO, L.S. and SCHUMAN, J.S., 2011. Optical Coherence Tomography: History, Current Status, and Laboratory Work. *Investigative Ophthalmology & Visual Science,* **52**, pp. 2425-2436.

GARDINER, S.K. and CRABB, D.P., 2002a. Examination of Different Pointwise Linear Regression Methods for Determining Visual Field Progression. *Investigative Ophthalmology & Visual Science,* **43**(5), pp. 1400-1407.

GARDINER, S.K. and CRABB, D.P., 2002b. Frequency of testing for detecting visual field progression. *British Journal of Ophthalmology,* **86**, pp. 560-564.

GARDINER, S.K. and JOHNSON, C.A., 2012. The effect of test variability on the structure-function relationship in early glaucoma. *Graefe's Archive for Clinical and Experimental Ophthalmology,* **250**, pp. 1851-1861.

GARDINER, S.K., DEMIREL, S., GORDON, M.O., KASS, M.A. and FOR THE OCULAR HYPERTENSITON TREATMENT STUDY GROUP, 2013. Seasonal Changes in Visual Field Sensitivity and Intraocular Pressure in the Ocular Hypertension Treatment Study. *Ophthalmology,* **120**(4), pp. 724-730.

GARDINER, S.K., SWANSON, W.H., GOREN, D., MANSBERGER, S.L. and DEMIREL, S., 2014. Assessment of the Reliability of Standard Automated Perimetry in Regions of Glaucomatous Damage. *Ophthalmology,* **121**(7), pp. 1359-1369.

GIANGIACOMO, A., GARWAY-HEATH, D.F. and CAPRIOLI, J., 2006. Diagnosing glaucoma progression: current practice and promising technologies. *Current Opinion in Ophthalmology,* **17**, pp. 153-162.

GILLESPIE, B.W., MUSCH, D.C., GUIRE, K.E., MILLS, R.P., LICHTER, P.R., JANZ, N.K., WREN, P.A. and THE CIGTS STUDY GROUP, 2003. The Collaborative Initial Glaucoma Treatment Study: Baseline Visual Field and Test-Retest Variability. *Investigative Ophthalmology & Visual Science,* **44**, pp. 2613-2620.

GLASS, E., SCHAUMBERGER, M. and LACHENMAYR, B.J.,1995. Simulations for FASTPAC and the 4-2 dB full-threshold strategy of the Humphrey Field Analyzer. *Investigative Ophthalmology & Visual Science,* **36**, pp. 1847-1854.

GLEN, F.C., CRABB, D.P. and GARWAY-HEATH, D.F., 2011. The direction of research into visual disability and quality of life in glaucoma. *BMC Ophthalmology,* **11**(19).

GLEN, F.C., CRABB, D.P., SMITH, N.D., BURTON, R. and GARWAY-HEATH, D.F., 2012. Do patients with glaucoma have difficulty recognising faces? *Investigative Ophthalmology & Visual Science,* **53**(7), pp. 3629-3637.

GLEN, F.C., SMITH, N.D. and CRABB, D.P., 2014. 'What types of visual field defects are hazardous for driving?' paper presented at *The Association for Research in Vision and Ophthalmology Annual Meeting*, Orlando FL, 6th May 2014.

GOLDBERG, I., CLEMENT, C.I., CHIANG, T.H., WALT, J.G., LEE, L.J., GRAHAM, S. and HEALEY, P.R., 2009. Assessing quality of life in patients with glaucoma using the Glaucoma Quality of Life-15 (GQL-15) questionnaire. *Journal of Glaucoma*, **18**(1): pp. 6-12

GORDON, M.O., BEISER JULIA A., BRANDT, J.D., HEUER, D.K., HIGGINBOTHAM, E.J., JOHNSON, C.A., KELTNER, J.L., MILLER, J.P., PARRISH, R.K.I., WILSON, M.R. and KASS, M.A., 2002. The Ocular Hypertension Treatment Study - Baseline Factors That Predict the Onset of Primary Open-Angle Glaucoma. *Archives of Ophthalmology,* **120**(6), pp. 714-720.

GRANT, W.M. and BURKE, J.F.J., 1982. Why do some people go blind from glaucoma? *Ophthalmology,* **89**(9), pp. 991-998.

GRODUM, K., HEIJL, A. and BENGTSSON, B., 2004. Glaucoma and mortality. *Graefe's Archive for Clinical and Experimental Ophthalmology,* **242**, pp. 397-401.

GURWITZ, J.H., GLYNN, R.J., MONANE, M., EVERITT, D.E., GILDEN, D., SMITH, N. and AVORN, J., 1993. Treatment for Glaucoma: Adherence by the Elderly. *American Journal of Public Health,* **83**, pp. 711-716.

GUTIERREZ, P.R., WILSON, M.R., JOHNSON, C.A., GORDON, M.O., CIOFFI, G.A., RITCH, R., SHERWOOD, M.B., MENG, K. and MANGIONE, C.M., 1997. Influence of Glaucomatous Visual Field Loss on Health-Related Quality of Life. *Archives of Ophthalmology,* **115**, pp. 777-784.

HAAG-STREIT INTERNATIONAL, 2009. *OCTOPUS 900 & EyeSuite SOP - Standard operating procedure* [online] Available at: http://innz.no/files/manualer/octopus900_sop.pdf [Accessed 10th March 2012].

HARRIS, M.L. and JACOBS, N.A., 1994. Is the Esterman binocular field sensitive enough? R.P. MILLS and M. WALL, eds. In: *Perimetry Update 1994/1995 : proceedings of the XIth International Perimetric Society Meetin*g, 3-7 July 1994, Kugler Publications.

HATTENHAUER, M.G., JOHNSON, D.H., ING, H.H., HERMAN, D.C., HODGE, D.O., YAWN, B.P., BUTTERFIELD, L.C. and GREY, D.T., 1998. The probability of blindness from open-angle glaucoma. *Ophthalmology,* **105**(11), pp. 2099-2104.

HEEG, G.P., BLANKSMA, L.J., HARDUS, P.L. and JANSONIUS, N.M., 2005. The Groningen Longitudinal Glaucoma Study. I. Baseline sensitivity and specificity of the frequency doubling perimeter and the GDx nerve fibre analyser. *Acta Ophthalmologica Scandinavica*, **83**, pp. 46-52

HEIJL, A. and KRAKAU, C.E.T., 1975. An Automatic Static Perimeter Design and Pilot Study. *Acta Ophthalmologica,* **53**(3), pp. 293-310.

HEIJL, A. and KRAKAU, C.E.T., 1977. A Note on Fixation during Perimetry. *Acta Ophthalmologica Scandinavica,* **55**, pp. 854-861.

HEIJL, A., LINDGREN, G. and OLSSON, J., 1987. Normal Variability of Static Perimetric Threshold Values Across the Central Visual Field. *Archives of Ophthalmology,* **105**, pp. 1544-1549.

HEIJL, A., LINDGREN, G. and OLSSON, J., 1989. The Effect of Perimetric Experience in Normal Subjects. *Archives of Ophthalmology,* **107**, pp. 81-86.

HEIJL, A. and BENGTSSON, B., 1996. The Effect of Perimetric Experience in Patients With Glaucoma. *Archives of Ophthalmology,* **114**(1), pp. 19-22.

HEIJL, A., LESKE, M.C., BENGTSSON, B., HYMAN, L., BENGTSSON, B. and HUSSEIN, M., 2002. Reduction of Intraocular Pressure and Glaucoma Progression - Results From the Early Manifest Glaucoma Trial. *Archives of Ophthalmology,* **120**, pp. 1268-1279.

HEIJL, A., LESKE, M.C., BENGTSSON, B., BENGTSSON, B., HUSSEIN, M. and EMGT GROUP, 2003. Measuring visual field progression in the Early Manifest Glaucoma Trial. *Acta Ophthalmologica Scandinavica,* **81**, pp. 286-293.

HEIJL, A., BENGTSSON, B., CHAUHAN, B.C., LIEBERMAN, M.F., CUNLIFFE, I., HYMAN, L. and LESKE, M.C., 2008. A Comparison of Visual Field Progression Criteria of 3 Major Glaucoma Trials in Early Manifest Glaucoma Trial Patients. *American Academy of Ophthalmology,* **115**, pp. 1557-1565.

HEIJL, A., BENGTSSON, B., HYMAN, L., LESKE, M.C. and EMGT GROUP, 2009. Natural History of Open-Angle Glaucoma. *Ophthalmology,* **116**, pp. 2271-2276.

HEIJL, A., ASPBERG, J. and BENGTSSON, B., 2011. The effect of different criteria on the number of patients blind from open-angle glaucoma. *BMC Ophthalmology,* **11**(31).

HEIJL, A., BUCHHOLZ, P., NORRGREN, G. and BENGTSSON, B., 2012a. Rates of visual field progression in clinical glaucoma care. *Acta Ophthalmologica Scandinavica,* **91**(5), pp. 406-412.

HEIJL, A., PATELLA, V.M. and BENGTSSON, B., 2012b. *Effective Perimetry.* Fourth Edition. Dublin CA, United States: Carl Zeiss Meditec.

HEIJL, A., 2013. The times they are a-changin': time to change glaucoma management. *Acta Ophthalmologica Scandinavica,* **91**, pp. 92-99.

HENSON, D.B., EVANS, J., CHAUHAN, B.C. and LANE, C., 1996. Influence of Fixation Accuracy on Threshold Variability in Patients With Open Angle Glaucoma. *Investigative Ophthalmology & Visual Science*, 37, pp. 444-450

HENSON, D.B., ARTES, P.H. and CHAUHAN, B.C., 1999. Diffuse Loss of Sensitivity in Early Glaucoma. *Investigative Ophthalmology & Visual Science*, **40**, pp. 3147-3151

HENSON, D.B., 2000. *Visual Fields.* Second Edition. Oxford: Butterworth-Heinemann.

HENSON, D.B., CHAUDRY, S., ARTES, P.H., FARAGHER, E.B. and ANSONS, A., 2000. Response Variability in the Visual Field: Comparison of Optic Neuritis, Glaucoma, Ocular Hypertension and Normal Eyes. *Investigative Ophthalmology & Visual Science,* **4**1, pp. 417-421.

HENSON, D.B. and THAMPY, R., 2005. Preventing blindness from glaucoma - Better screening with existing tests should be the priority. *British Medical Journal,* **331**, pp. 120-121.

HENSON, D.B. and SHAMBHU, S., 2006. Relative Risk of Progressive Glaucomatous Visual Field Loss in Patients Enrolled and Not Enrolled in a Prospective Longitudinal Study. *Archives of Ophthalmology,* **124**, pp. 1405-1408.

HENSON, D.B. and EMUH, T., 2010. Monitoring Vigilance during Perimetry by Using Pupillography. *Investigative Ophthalmology & Visual Science,* **51**, pp. 3540-3543.

HIRNEISS, C., VOGEL, M., KAMPIK, A., NEUBAUER, A.S. and KERNT, M., 2011. Measurement of glaucoma-specific functionality with the GQL-15 and correlation with parameters of visual function. *Ophthalmologe*. **108**(10), pp. 939-46.

HODAPP, E., PARRISH, R.K.I. and ANDERSON, D.R., 1993. *Clinical decisions in glaucoma.* St Louis: The CV Mosby Co.

JAMPEL, H.D., 2001. Glaucoma Patients' Assessment of their Visual Function and Quality of Life. *Transactions of the American Ophthalmological Society,* **99**, pp. 301-317.

JAMPEL, H.D., FRIEDMAN, D.S., QUIGLEY, H.A. and MILLER, R., 2002a. Correlation of the binocular visual field with patient assessment of vision. *Investigative Ophthalmology & Visual Science,* **43**, pp. 1059-1067.

JAMPEL, H.D., SCHWARTZ, A., POLLACK, I., ABRAMS, D., WEISS, H. and MILLER, R., 2002b. Glaucoma patients' assessment of their visual function and quality of life. *Journal of Glaucoma,* **11**, pp. 154-163.b

JANSONIUS, N.M., 2006. Towards an optimal perimetric strategy for progression detection in glaucoma: from fixed-space to adaptive inter-test intervals. *Graefe's Archive for Clinical and Experimental Ophthalmology,* **24**4(3), pp. 390-393.

JANSONIUS, N.M., 2007. Progression detection in glaucoma can be made more efficient by using a variable interval between successive visual field tests. *Graefe's Archive for Clinical and Experimental Ophthalmology,* **24**5(11), pp. 1647-1651.

JANSONIUS, N.M., 2010. On the accuracy of measuring rates of visual field change in glaucoma. *British Journal of Ophthalmology,* **94**, pp. 1404-1405.

JANZ, N.K., WREN, P.A., LICHTER, P.R., MUSCH, D.C., GILLESPIE, B.W. and GUIRE, K.E., 2001. Quality of life in newly diagnosed glaucoma patients: The Collaborative Initial Glaucoma Treatment Study. *Ophthalmology,* **108,** pp. 887-897.

JIANG, X., VARMA, R., WU, S., TORRES, M., AZEN, S.P., FRANCIS, B.A., CHOPRA, V. and NGUYEN, B.B., 2012. Baseline Risk Factors that Predict the Development of Open-Angle Glaucoma in a Population. *American Academy of Ophthalmology,* **119**, pp. 2245-2253.

JOHNSON, C.A. and KELTNER, J.L., 1983. Incidence of visual field loss in 20,000 eyes and its relationship to driving performance. *Archives of Ophthalmology,* **101**, pp. 371-375.

JONAS, J.B., NGUYEN, X.N., GUSEK, G.C. and NAUMANN, G.O., 1989. Parapapillary chorioretinal atrophy in normal and glaucoma eyes. I. Morphometric data. *Investigative Ophthalmology & Visual Science,* **30(5)**, pp. 908-918.

JUNOY MONTOLIO, F.G., WESSELINK, C., GORDIJN, M. and JANSONIUS, N.M., 2012. Factors that influence standard automated perimetry test results in glaucoma: test reliability, technician experience, time of day and season. *Investigative Ophthalmology & Visual Science,* **53**(11), pp. 7010-7017.

KHAWAJA, A.P., CHAN, M.P.Y., HAYAT, S., BROADWAY, D.C., LUBEN, R., GARWAY-HEATH, D.F., SHERWIN, J.C., YIP, J.L.Y., DALZELL, N., WAREHAM, N.J.,

KHAW, K. and FOSTER, P.J., 2013. The EPIC-Norfolk Eye Study: rationale, methods and cross-sectional analysis of visual impairment in a population-based cohort. *BMJ Open,* **3**, pp. e002684.

KHAWAJA, A.P., CRABB, D.P. and JANSONIUS, N.M., 2013. The Role of Ocular Perfusion Pressure in Glaucoma Cannot Be Studied With Multivariate Regression Analysis Applied to Surrogates. *Investigative Ophthalmology & Visual Science,* **54**, pp. 4619-4620.

KIM, C.S., SEONG, G.J., LEE, N.H., SONG, K.C., NAMIL STUDY GROUP and KOREAN GLAUCOMA SOCIETY, 2011. Prevalence of primary open-angle glaucoma in central South Korea: the Namil study. *Ophthalmology,* **118**(6), pp. 1024-1030.

KIM, J.M., KYUNG, H., AZARBOD, P., LEE, J.M. and CAPRIOLI, J., 2014. Disc haemorrhage is associated with the fast component, but not the slow component of visual field decay rate in glaucoma. *British Journal of Ophthalmology*, **In Press**

KOBELT, G., JONSSON, B., BERGSTROM, A., CHEN, E., LINDEN, C. and ALM, A., 2006. Cost-effectiveness analysis in glaucoma: what drives utility? Results from a pilot study in Sweden. *Acta Ophthalmologica Scandinavica,* **84**, pp. 363-371.

KONO, Y., ZANGWILL, L.M., SAMPLE, P.A., JONAS, J.B., EMDADI, A., GUPTA, N. and WEINREB, R.N., 1999. Relationship between parapapillary atrophy and visual field abnormality in primary open-angle glaucoma. *American Journal of Ophthalmology,* **127**(6), pp. 674-680.

KOOIJMAN, A.C., BROUWER, W.H., COECKELBERGH, T.R.M., TANT, M.L.M., CORNELISSEN, F.W., BREDEWOUD, R.A. and MELIS-DANKERS, B.J.M., 2004. Compensatory viewing training improves practical fitness to drive of subjects with impaired vision. *Visual Impairment Research,* **6**(1), pp. 1-27.

KOTECHA, A., SPRATT, A. and VISWANATHAN, A.C., 2008. Visual function and fitness to drive. *British Medical Bulletin,* **87**, pp. 163-174.

KOTECHA, A., O'LEARY, N., MELMOTH, D., GRANT, S. and CRABB, D.P., 2009. The Functional Consequences of Glaucoma for Eye-Hand Coordination. *Investigative Ophthalmology & Visual Science,* **50**, pp. 203-213.

KOTECHA, A., FERNANDES, S., BUNCE, C. and FRANKS, W.A., 2012a. Avoidable sight loss from glaucoma: is it unavoidable? *British Journal of Ophthalmology,* **96**, pp. 816-820.

KOTECHA, A., RICHARDSON, G., CHOPRA, V., FAHY, R.T., GARWAY-HEATH, D.F. and RUBIN, G.S., 2012b. Balance control in glaucoma. *Investigative Ophthalmology & Visual Science,* **53**(12), pp. 7795-7801.

KWON, Y.H., KIM, C., ZIMMERMAN, M.B., ALWARD, W.L.M. and HAYREH, S.S., 2001. Rate of visual field loss and long-term visual outcome in primary open-angle glaucoma. *American Journal of Ophthalmology,* **132**(1), pp. 47-56.

LABIRIS, G., KATSANOS, A., FANARIOTIS, M., ZACHARAKI, F., CHATZOULIS, D. and KOZOBOLIS, V.P., 2010. Vision-specific quality of life in Greek glaucoma patients. *Journal of Glaucoma,* **19**(1), pp. 39-43.

LAMING, D. and LAMING, J., 1992. F. Hegelmaier: on memory for the length of a line. *Psychological Research,* **54**(4), pp. 233-239.

LEE, B.L., GUTIERREZ, P.R., GORDON, M.O., WILSON, M.R., CIOFFI, G.A., RITCH, R., SHERWOOD, M.B. and MANGIONE, C.M., 1998. The Glaucoma Symptom Scale - A brief index of glaucoma-specific symptoms. *Archives of Ophthalmology,* **116**, pp. 861-866.

LEE, J.M., CAPRIOLI, J., NOURI-MAHDAVI, K., AFIFI, A.A., MORALES, E., RAMANATHAN, M., YU, F. and COLEMAN, A.L., 2014. Baseline Prognostic Factors Predict Rapid Visual Field Deterioration in Glaucoma. *Investigative Ophthalmology & Visual Science,* **55**, pp. 2228-2236.

LESKE, M.C., CONNELL, A.M.S., WU, S., NEMESURE, B., LI, X., SCHACHAT, A., HENNIS, A. and HENNIS, A., 2001. Incidence of Open-Angle Glaucoma - The Barbados Eye Studies. *Archives of Ophthalmology,* **119**, pp. 89-95.

LESKE, M.C., HEIJL, A., HUSSEIN, M., BENGTSSON, B., HYMAN, L. and KOMAROFF, E., 2003. Factors for Glaucoma Progression and the Effect of Treatment. *Archives of Ophthalmology,* **121**, pp. 48-56.

LESKE, M.C., HEIJL, A., HYMAN, L., BENGTSSON, B., DONG, L.M. and YANG, Z.Y., 2007. Predictors of Long-term Progression in the Early Manifest Glaucoma Trial. *Ophthalmology,* **114**, pp. 1965-1972.

LESKE, M.C., WU, S.M.A., HENNIS, A., HONKANEN, R. and NEMESURE, B., 2008. Risk Factors for Incident Open-angle Glaucoma - The Barbados Eye Studies. *American Academy of Ophthalmology,* **115**, pp. 85-93.

LISBOA, R., CHUN, Y.S., ZANGWILL, L.M., WEINREB, R.N., ROSEN, P.N., LIEBMANN, J.M., GIRKIN, C.A. and MEDEIROS, F.A., 2013. Association Between Rates of Binocular Visual Field Loss and Vision-Related Quality of Life in Patients With Glaucoma. *Journal of the American Medical Association Ophthalmology,* **131**(4), pp. 486-494.

MALIK, R., SWANSON, W.H. and GARWAY-HEATH, D.F., 2006. Development and evaluation of a linear staircase strategy for the measurement of perimetric sensitivity. *Vision Research,* **46**, pp. 2956-2967.

MALIK, R., BAKER, H., RUSSELL, R.A. and CRABB, D.P., 2013. A survey of attitudes of glaucoma subspecialists in England and Wales to visual field test intervals in relation to NICE guidelines. *BMJ Open,* **3**(5).

MANGIONE, C.M., BERRY, S., SPRITZER, K., JANZ, N.K., KLEIN, R., OWSLEY, C. and LEE, P.P., 1998. Identifying the Content Area for the 51-Item National Eye Institute Visual Function Questionnaire - Results From Focus Groups With Visually Impaired Persons. *Archives of Ophthalmology,* **116**, pp. 227-233.

MANSBERGER, S.L. and CIOFFI, G.A., 2006. The Probability of Glaucoma From Ocular Hypertension Determined by Ophthalmologists in Comparison to a Risk Calculator. *Journal of Glaucoma,* **15**, pp. 426-431.

MARCUS, M.W., DE VRIES, M.M., JUNOY MONTOLIO, F.G. and JANSONIUS, N.M., 2011. Myopia as a Risk Factor for Open-Angle Glaucoma: A Systematic Review and Meta-Analysis. *Ophthalmology,* **118**(10), pp. 1989-1994.

MAYCOCK, G. and LOCKWOOD, C.R., 1993. The accident liability of British car drivers. *Transport Reviews: A Transnational Transdisciplinary Journal,* **13**(3), pp. 231-245.

MCCLOSKEY, L.W., KOEPSELL, T., WOLF, M.E. and BUCHNER, D.M., 1994. Motor vehicle collision injuries and sensory impairments of older drivers. *Age and Ageing,* **23**, pp. 267-273.

MCKEAN-COWDIN, R., WANG, Y., WU, J., AZEN, S.P., VARMA, R. and LOS ANGELES LATINO EYE STUDY GROUP, 2008. Impact of Visual Field Loss on Health-Related Quality of Life in Glaucoma - The Los Angeles Latino Eye Study. *Ophthalmology,* **115**, pp. 941-948.

MCNAUGHT, A.I., CRABB, D.P., FITZKE, F.W. and HITCHINGS, R.A., 1995. Modelling series of visual fields to detect progression in normal-tension glaucoma. *Graefe's Archive for Clinical and Experimental Ophthalmology,* **233**, pp. 750-755.

MEDEIROS, F.A., SAMPLE, P.A., ZANGWILL, L.M., BOWD, C., AIHARA, M. and WEINREB, R.N., 2003. Corneal Thickness as a Risk Factor for Visual Field Loss in Patients With Preperimetric Glaucomatous Optic Neuropathy. *American Journal of Ophthalmology*, **136**, pp. 805-813

MEDEIROS, F.A., LEITE, M.T., ZANGWILL, L.M. and WEINREB, R.N., 2011. Combining Structural and Functional Measurements to Improve Detection of Glaucoma Progression using Bayesian Hierarchical Models. *Investigative Ophthalmology & Visual Science,* **52**, pp. 5794-5803.

MEDEIROS, F.A., ZANGWILL, L.M., MANSOURI, K., LISBOA, R., TAFRESHI, A. and WEINREB, R.N., 2012. Incorporating Risk Factors to Improve the Assessment of Rates of Glaucomatous Progression. *Investigative Ophthalmology & Visual Science,* **53**(4), pp. 2199-2207.

MEMBREY, W.L., POINOOSAUMY, D.P., BUNCE, C., FITZKE, F.W. and HITCHINGS, R.A., 2000. Comparison of visual field progression in patients with normal pressure glaucoma between eyes with and without visual field loss that threatens fixation. *British Journal of Ophthalmology,* **84**, pp. 1154-1158.

MILLS, R.P., BUDENZ, D.L., LEE, P.P., NOECKER, R.J., WALT, J.G., SIEGARTEL, L.R., EVANS, S.J. and DOYLE, J.J., 2006. Categorising the stage of glaucoma from pre-diagnosis to end-stage disease. *American Journal of Ophthalmology,* **141**, pp. 24-30.

MITCHELL, P., SMITH, W., CHEY, T. and HEALEY, P.R., 1997. Open-angle Glaucoma and Diabetes: The Blue Mountains Eye Study, Australia. *Ophthalmology,* **104**, pp. 712-718.

MOJON, D.S., HESS, C.W., GOLDBLUM, D., BOHNKE, M., KORNER, F. and MATHIS, J., 2000. Primary open-angle glaucoma is associated with sleep apnea syndrome. *Ophthalmologica,* **214**, pp. 115-118.

MUKESH, B.N., MCCARTY, C.A., RAIT, J.L. and TAYLOR, H.R., 2002. Five-Year Incidence of Open-angle Glaucoma - The Visual Impairment Project. *Ophthalmology,* **109**, pp. 1047-1051.

MURATA, H., HIRASAWA, H., AOYAMA, Y., SUGISAKI, K., ARAIE, M., MAYAMA, C., AIHARA, M. and ASAOKA, R., 2013. Identifying Areas of the VIsual Field Important for Quality of Life in Patients with Glaucoma. *PLoS ONE,* **8**(3), pp. 1-7.

MUSCH, D.C., LICHTER, P.R., GUIRE, K.E., STANDARDI, C.L. and THE CIGTS STUDY GROUP, 1999. The Collaborative Initial Glaucoma Treatment Study - Study Design, Methods, and Baseline Characteristics of Enrolled Patients. *Ophthalmology,* **106**(4), pp. 653-662.

MUSCH, D.C., GILLESPIE, B.W., MOTYKA, B.M., NIZIOL, L.M., MILLS, R.P. and LICHTER, P.R., 2005. Converting to SITA-Standard from Full-Threshold Visual Field Testing in the Follow-up Phase of a Clinical Trial. *Investigative Ophthalmology & Visual Science,* **46**, pp. 2755-2759.

NATIONAL EYE INSTITUTE, 2010, *Glaucoma, Open-angle* [online]. Available at: http://www.nei.nih.gov/eyedata/glaucoma.asp [Accessed 24th June 2014].

NATIONAL INSTITUTE FOR HEALTH AND CLINICAL EXCELLENCE, 2009, NICE Clinical Guidelines - Glaucoma: Diagnosis and management of chronic open angle glaucoma and ocular hypertension [online] Available at: http://publications.nice.org.uk/glaucoma-quality-standard-qs7 [Accessed 17th June 2014].

NELSON, P., ASPINALL, P., A. and O'BRIEN, C., 1999. Patients' perception of visual impairment in glaucoma: a pilot study. *British Journal of Ophthalmology,* **83**, pp. 546-552.

NELSON-QUIGG, J., CELLO, K. and JOHNSON, C.A., 2000. Predicting Binocular Visual Field Sensitivity from Monocular Visual Field Results. *Investigative Ophthalmology & Visual Science,* **41**, pp. 2212-2221.

NOE, G., FERRARO, J., LAMOUREUX, E., RAIT, J. and KEEFFE, J.E., 2003. Associations between glaucomatous visual field loss and participation in activities of daily living. *Clinical and Experimental Ophthalmology,* **31**, pp. 482-486.

NORDMANN, J.P., BRION, F., HAMARD, P. and MOUTON-CHOPIN, D., 1998. Evaluation of the Humphrey perimetry programs SITA Standard and SITA Fast in

normal probands and patients with glaucoma. *French Journal of Ophthalmology*. **21**(8), pp. 549-554.

NOURI-MAHDAVI, K., BRIGATTI, L., WEITZMAN, M. and CAPRIOLI, J., 2007. Comparison of methods to detect visual field progression in glaucoma. *Archives of Ophthalmology,* **125**(9), pp. 1176-1181.

NOURI-MAHDAVI, K. and CAPRIOLI, J., 2014. Measuring rates of structural and functional change in glaucoma. *British Journal of Ophthalmology*, Online First

OCULAR HYPERTENSION TREATMENT STUDY GROUP and EUROPEAN GLAUCOMA PREVENTION STUDY GROUP, 2007. Validated Prediction Model for the Development of Primary Open-Angle Glaucoma in Individuals with Ocular Hypertension. *Ophthalmology,* **114**, pp. 10-19.

OFFICE OF NATIONAL STATISTICS, 2012. *Ethnicity and National Identity in England and Wales 2011,* [online] Available at : http://www.ons.gov.uk/ons/dcp171776_290558.pdf [Accessed 11th March 2013].

OFFICE OF NATIONAL STATISTICS, 2011. *England, Interim Life Tables, 1980-82 to 2008-10*, [online] Available at: http://ons.gov.uk/ons/rel/lifetables/interim-life-tables/2008-2010/rft-ilt-eng-2008-10.xls [Accessed 31st August 2012].

O'LEARY, N., CHAUHAN, B.C. and ARTES, P.H., 2012. Visual field progression in glaucoma: estimating the overall significance of deterioration with permutation analyses of pointwise linear regression (PoPLR). *Investigative Ophthalmology & Visual Science,* **53**(11), pp. 6776-6784.

OLIVER, J.E., HATTENHAUER, M.G., HERMAN, D., HODGE, D.O., KENNEDY, R., FANG-YEN, M. and JOHNSON, D.H., 2002. Blindness and glaucoma: a comparison of patients progressing to blindness from glaucoma with patients maintaining vision. *American Journal of Ophthalmology,* **133**(6), pp. 764-772.

OLKIN, I. and FINN, J.D., 1995. Correlations Redux. *Psychological Bulletin,* **118**(1), pp. 155-164.

OWEN, V.M.F., STROUTHIDIS, N.G., GARWAY-HEATH, D.F. and CRABB, D.P., 2006. Measurement Variability in Heidelberg Retina Tomograph Imaging of Neuroretinal Rim Area. *Investigative Ophthalmology & Visual Science,* **47**(12), pp. 5322-5330.

OWEN, V.M.F., CRABB, D.P., WHITE, E.T., VISWANATHAN, A.C., GARWAY-HEATH, D.F. and HITCHINGS, R.A., 2008. Glaucoma and Fitness to Drive: Using Binocular Visual Fields to Predict a Milestone to Blindness. *Investigative Ophthalmology & Visual Science,* **49**, pp. 2449-2455.

OWSLEY, C., MCGWIN, G. and BALL, K., 1998a. Vision impairment, eye disease and injurious motor vehicle crashes in the elderly. *Ophthalmic Epidemiology,* 5, pp. 101-113.

OWSLEY, C., BALL, K., MCGWIN, G., SLOANE, M.E., ROENKER, D.L., WHITE, M.F. and OVERLEY, T., 1998b. Visual Processing Impairment and Risk of Motor Vehicle Crash Among Older Adults. *Journal of the American Medical Association,* **279**, pp. 1083-1088.

PARRISH, R.K.I., GEDDE, S.J., SCOTT, I.U., FEUER, W.J., SCHIFFMAN, J.C., MANGIONE, C.M. and MONTENEGRO-PINIELLA, A., 1997. Visual function and quality of life among patients with glaucoma. *Archives of Ophthalmology,* **115**, pp. 1447-1455.

PATHAK, M., DEMIREL, S. and GARDINER, S.K., 2013. Nonlinear, Multilevel Mixed-Effects Approach for Modeling Longitudinal Standard Automated Perimetry Data in Glaucoma. *Investigative Ophthalmology & Visual Science,* **54**, pp. 5505-5513.

PERDICCHI, A., IESTER, M., SCUDERI, G., AMODEO, S., MEDORI, E.M. and RECUPERO, S.M., 2007. Visual field damage and progression in glaucomatous myopic eyes. *European Journal of Ophthalmology,* **17**(4), pp. 534-537.

PIERRE-FILHO, P.deT., SCHIMITI, R.B., DE VASCONCELLOS, J.P. and COSTA, V.P., 2006. Sensitivity and specificity of frequency-doubling technology, tendency-oriented perimetry, SITA Standard and SITA Fast perimetry in perimetrically inexperienced individuals. *Acta Ophthalmologica Scandinavica,* **84**, pp. 345-350.

PRATA, T.S., DE MORAES, C.G.V., TENG, C.C., TELLO, C., RITCH, R. and LIEBMANN, J.M., 2010. Factors Affecting Rates of Visual Field Progression in Glaucoma Patients with Optic Disc Hemorrhage. *Ophthalmology,* **117**, pp. 24-29.

PUVANACHANDRA, N., KANG, C.Y., KIRWAN, J.F. and JEFFREY, M.N., 2008. How good are we at advising appropriate patients with glaucoma to inform the DVLA? A closed audit loop. *Ophthalmic and Physiological Optics,* **28**, pp. 313-316.

QUIGLEY, H.A. and BROMAN, A.T., 2006. The number of people with glaucoma worldwide in 2010 and 2020. *British Journal of Ophthalmology,* **90**, pp. 262-267.

R DEVELOPMENT CORE TEAM, 2014. *R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.* Available: http://www.R-project.org/.

RACETTE, L. and CASSON, E.J., 2005. The Impact of Visual Field Loss on Driving Performance: Evidence From On-Road Driving Assessments. *Optometry and Vision Science,* **82**(8), pp. 668-674.

RADCLIFFE, N.M., CHUNG, E., DEMETRIADES, A.M. and CHRISTOS, P.J., 2014. Structural glaucomatous progression before and after occurrence of an optic disc haemorrhage. *British Journal of Ophthalmology*, **In Press**

RAMULU, P.Y., 2009. Glaucoma and disability: which tasks are affected, and at what stage of disease? *Current Opinion in Ophthalmology,* **20**(2), pp. 92-98.

RAMULU, P.Y., WEST, S.K., MUNOZ, B., JAMPEL, H.D. and FRIEDMAN, D.S., 2009a. Glaucoma and Reading Speed - The Salisbury Eye Evaluation Project. *Archives of Ophthalmology,* **127**(1), pp. 82-87.

RAMULU, P.Y., WEST, S.K., MUNOZ, B., JAMPEL, H.D. and FRIEDMAN, D.S., 2009b. Driving Cessation and Driving Limitation in Glaucoma - The Salisbury Eye Evaluation Project. *Ophthalmology,* **116**, pp. 1846-1853.

RAMULU, P.Y., SWENOR, B.K., JEFFERYS, J.L., FRIEDMAN, D.S. and RUBIN, G.S., 2012. Difficulty with Out-loud and Silent Reading in Glaucoma. *Investigative Ophthalmology & Visual Science,* **54**(1), pp. 666-672.

REIS, A.S.C., ARTES, P.H., BELLIVEAU, A.C., LEBLANC, R.P., SHUBA, L.M., CHAUHAN, B.C. and NICOLELA, M.T., 2012. Rates of Change in the Visual Field and Optic Disc in Patients with Distinct Patterns of Glaucomatous Optic Disc Damage. *Ophthalmology*, **119**(2), pp. 294-303

RICHMAN, J., LORENZANA, L.L., LANKARANIAN, D., DUGAR, J., MAYER, J.R., WIZOV, S.S. and SPAETH, G.L., 2010. Relationships in Glaucoma Patients Between Standard Vision Tests, Quality of Life, and Ability to Perform Daily Activities. *Ophthalmic Epidemiology,* **17**(3), pp. 144-151.

ROBERTS, K.F., HAYMES, S.A., LEBLANC, R.P., NICOLELA, M.T., CHAUHAN, B.C. and ARTES, P.H., 2005. 'Contrast Sensitivity, Visual Acuity, Reading Speed and Macular Visual Field Damage in Glaucoma.' paper presented at *The Association for Research in Vision and Ophthalmology Annual Meeting*, Fort Lauderdale FL.

GLEN, F.C., SMITH, N.D. and CRABB, D.P., 2014. 'What types of visual field defects are hazardous for driving?' paper presented at *The Association for Research in Vision and Ophthalmology Annual Meeting*, Orlando FL, 6th May 2014.

ROBIN, X., TURCK, N., HAINARD, A., TIBERTI, N., LISACEK, F., SANCHEZ, J. and MLLER, M., 2011. pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics*, **12**:77.

ROYAL COLLEGE OF OPHTHALMOLOGISTS, 2013. *Visual Standards for Driving*. [online] Available at: http://www.rcophth.ac.uk/pdf [Accessed 16th April 2014].

RUSSELL, R.A. and CRABB, D.P., 2011. On Alternative Methods for Measuring Visual Field Decay: Tobit Linear Regression. *Investigative Ophthalmology & Visual Science,* **52**(13), pp. 9539-9540.

RUSSELL, R., A., CRABB, D.P., MALIK, R. and GARWAY-HEATH, D.F., 2012a. The Relationship between Variability and Sensitivity in Large-Scale Longitudinal Visual Field Data. *Investigative Ophthalmology & Visual Science,* **53**(10), pp. 5985-5990.

RUSSELL, R.A., MALIK, R., CHAUHAN, B.C., CRABB, D.P. and GARWAY-HEATH, D.F., 2012b. Improved Estimates of Visual Field Progression Using Bayesian

Linear Regression to Integrate Structural Information in Patients with Ocular Hypertension. *Investigative Ophthalmology & Visual Science,* **53**, pp. 2760-2769.

RUSSELL, R.A., GARWAY-HEATH, D.F. and CRABB, D.P., 2013. New Insights into Measurement Variability in Glaucomatous Visual Fields from Computer Modelling. *PLoS One,* **8**(12): e83595.

SAUNDERS, L.J., RUSSELL, R.A. and CRABB, D.P., 2012a. Practical landmarks for visual field disability in glaucoma. *British Journal of Ophthalmology,* **96**(9), pp. 1185-1189.

SAUNDERS, L.J., RUSSELL, R.A. and CRABB, D.P., 2012b. The Coefficient of Determination: What Determines a Useful R^2 Statistic? *Investigative Ophthalmology & Visual Science,* **53**(11), pp. 6830-6832.

SAUNDERS, L.J., RUSSELL, R.A. and CRABB, D.P., 2013. Detecting Visual Field Progression in Glaucoma – Using the Right Tools for the Job, *European Ophthalmic Review,***7**(1), pp. 20-26.

SAUNDERS, L.J., RUSSELL, R.A., KIRWAN, J.F., MCNAUGHT, A.I. and CRABB, D.P., 2014. Examining Visual Field Loss in Patients in Glaucoma Clinics During Their Predicted Remaining Lifetime. *Investigative Ophthalmology & Visual Science,* **55**, pp. 102-109.

SAUNDERS, L.J., RUSSELL, R.A. and CRABB, D.P., 2015. Measurement Precision in a Series of Visual Fields Acquired by the Standard and Fast Versions of the Swedish Interactive Thresholding Algorithm: Analysis of Large-Scale Data From Clinics, *JAMA Ophthalmology,* **133**(1): 74-80.

SHAO, X., FENERTY, C. and HENSON, D.B., 2011. The Validity of Reliability Measure in Threshold Perimetry. *Clinical & Experimental Ophthalmology*, **2**:117

SHAW, M.E., 2005. Increasing compliance with glaucoma therapy: "so, convince me I have something wrong with my eyes". *Insight,* **30**(3), pp. 7-9.

SIEGNER, S.W. and NETLAND, P.A., 1996. Optic Disc Hemorrhages and Progression of Glaucoma. *Ophthalmology*, **103**(7), pp. 1014-1024

SINCLAIR, A., HINDS, A and SANDERS, R., 2004. Ten years of glaucoma blindness in Fife 1990-99 and the implications for ophthalmology, optometry and rehabilitation services. *Ophthalmic and Physiological Optics*, **24**(4), pp. 313-318

SKALICKY, S.E. and GOLDBERG, I., 2008. Depression and Quality of Life in Patients With Glaucoma: A Cross-sectional Analysis Using the Geriatric Depression Scale-15, Assessment of Function Related to Vision, and the Glaucoma Quality of Life-15. *Journal of Glaucoma,* **17**(7), pp. 546-551.

SKALICKY, S.E. and GOLDBERG, I., 2010. Are We Ready to Assess Quality of Life Routinely in our Glaucoma Patients? *Bulletin of the Belgian Society of Ophthalmology,* **315**, pp. 5-7.

SMITH, N.D., CRABB, D.P. and GARWAY-HEATH, D.F., 2011. An exploratory study of visual search performance in glaucoma. *Ophthalmic and Physiological Optics,* **31**, pp. 225-232.

SMITH, N.D., GLEN, F.C. and CRABB, D.P., 2012. Eye movements during visual search in patients with glaucoma. *BMC Ophthalmology*, **12**:45.

SMITH, S.D., KATZ, J. and QUIGLEY, H.A., 1996. Analysis of Progressive Change in Automated Visual Fields in Glaucoma. *Investigative Ophthalmology & Visual Science,* **37**, pp. 1419-1428.

SOBOTTKA VENTURA, A.C., BOHNKE, M., MOJON, D.S., 2001. Central corneal thickness measurements in patients with normal tension glaucoma, primary open angle glaucoma, pseudoexfoliation glaucoma, or ocular hypertension. *British Journal of Ophthalmology*, **85**, pp. 792-795

SOHN, S.W., SONG, J.S. and KEE, C., 2010. Influence of the Extent of Myopia on the Progression of Normal Tension Glaucoma. *American Journal of Ophthalmology,* **149**, pp. 831-838.

SOMMER, A., TIELSCH, J.M., KATZ, J., QUIGLEY, H.A., GOTTSCH, J.D., JAVITT, J. and SINGH, K., 1991. Relationship Between Intraocular Pressure and Primary Open Angle Glaucoma Among White and Black Americans - The Baltimore Eye Survey. *Archives of Ophthalmology,* **109**, pp. 1090-1095.

SPAETH, G.L.M., WALT, J.G. and KEENER, J., 2006. Evaluation of Quality of Life for Patients With Glaucoma. *American Journal of Ophthalmology,* **141**, pp. 3-14.

SPONSEL, W.E., GROTH, S.L., SATSANGI, N., MADDESS, T.L. and REILLY, M.A., 2014. Refined Data Analysis Provides Clinical Evidence for Central Nervous System Control of Chronic Glaucomatous Neurodegeneration. *Translational Vision Science and Technology,* **3**, pp. 3.

STROUTHIDIS, N.G. and GARWAY-HEATH, D.F., 2008. New developments in Heidelberg retina tomograph for glaucoma. *Current Opinion in Ophthalmology,* **19**, pp. 141-148.

SUSANNA JR., R. and VESSANI, R.M., 2009. Staging Glaucoma Patient: Why and How? *The Open Ophthalmology Journal,* **3**, pp. 59-64.

TABRETT, D.R. and LATHAN, K., 2012. Important areas of the central binocular visual field for daily functioning in the visually impaired. *Ophthalmic and Physiological Optics,* **32**, pp. 156-163.

TANNA, A.P., BANDI, J.R., BUDENZ, D.L., FEUER, W.J., FELDMAN, R.M., HERNDON, L.W., RHEE, D.J. and WHITESIDE-DE VOS, J., 2011. Interobserver Agreement and Intraobserver Reproducibility of the Subjective Determination of Glaucomatous Visual Field Progression. *Ophthalmology,* **118**, pp. 60-65.

TENG, C.C., DE MORAES, C.G.V., PRATA, T.S., TELLO, C., RITCH, R. and LIEBMANN, J.M., 2010. β-Zone Parapapillary Atrophy and the Velocity of Glaucoma Progression. *Ophthalmology,* **117**, pp. 909-915.

TEZEL, G., DORR, D., KOLKER, A.E., WAX, M.B. and KASS, M.A., 2000. Concordance of parapapillary chorioretinal atrophy in ocular hypertension with visual field defects that accompany glaucoma development. *Ophthalmology,* **107**(6), pp. 1194-1199.

TURALBA, A.V. and GROSSKREUTZ, C., 2010. A Review of Current Technology Used in Evaluating Visual Function in Glaucoma. *Seminars in Ophthalmology,* **25**(5-6), pp. 309-316.

TURPIN, A. and MCKENDRICK, A.M., 2005. Observed-based rather than population-based confidence limits for determining probability of change in visual fields. *Vision Research,* **45**, pp. 3277-3289.

TURPIN, A. and MCKENDRICK, A.M., 2011. What Reduction in Standard Automated Perimetry Variability Would Improve the Detection of Visual Field Progression? *Investigative Ophthalmology & Visual Science,* **52**, pp. 3237-3245.

TUULONEN, A., 2013. Challenges of glaucoma care - high volume, high quality, low cost. *Acta Ophthalmologica Scandinavica,* **91**, pp. 3-5.

US SOCIAL SECURITY ADMINISTRATION, 2011. *Disability Evaluation Under Social Security* [online] Available at: http://www.ssa.gov/disability/professionals/bluebook/2.00-SpecialSensesandSpeech-Adult.htm#203 [Accessed 31st January 2013].

VAN GESTEL, A., WEBERS, C.A.B., BECKERS, H.J.M., VAN DONGEN, M. C. J. M., SEVERENS, J.L., HENDRIKSE, F. and SCHOUTEN, JAN S. A. G., 2010. The relationship between visual field loss in glaucoma and health-related quality-of-life. *Eye,* **24**, pp. 1759-1769.

VAN LANDINGHAM, S.W., HOCHBERG, C., MASSOF, R.W., CHAN, E.S., FRIEDMAN, D.S. and RAMULU, P.Y., 2013. Driving patterns in older adults with glaucoma. *BMC Ophthalmology,* **13:**4.

VEGA, R.P., VAN LEEUWEN, P.M., RENDON VELEZ, E., LEMIJ, H.G. and DE WINTER, J.C.F., 2013. Obstacle Avoidance, Visual Detection Performance, and Eye-Scanning Behavior of Glaucoma Patients in a Driving Simulator: A Preliminary Study. *PLOS ONE,* **8**(10),.

VESTI, E., JOHNSON, C.A. and CHAUHAN, B.C., 2003. Comparison of Different Methods for Detecting Glaucomatous Visual Field Progression. *Investigative Ophthalmology & Visual Science,* **44**(9), pp. 3873-3879.

VISWANATHAN, A.C., MCNAUGHT, A.I., POINOOSAWMY, D., FONTANA, L., CRABB, D.P., FITZKE, F.W. and HITCHINGS, R.A., 1999. Severity and Stability of Glaucoma – Patient Perception Compared With Objective Measurement. *Archives of Ophthalmology*, **117**(4), pp. 450-454

VISWANATHAN, A.C., CRABB, D.P., MCNAUGHT, A.I., WESTCOTT, M.C., KAMAL, D., GARWAY-HEATH, D.F., FITZKE, F.W. and HITCHINGS, R.A., 2003. Interobserver agreement on visual field progression in glaucoma: a comparison of methods. *British Journal of Ophthalmology,* **87**, pp. 726-730.

WEINREB, R.N. and KAUFMAN, P.L., 2011. Glaucoma Research Community and FDA Look to the Future, II: NEI/FDA Glaucoma Clinical Trial Design and Endpoints Symposium: Measures of Structural Change and Visual Function. *Investigative Ophthalmology & Visual Science,* **52**(11), pp. 7842-7851.

WESSELINK, C., STOUTENBEEK, R. and JANSONIUS, N.M., 2011. Incorporating life expectancy in glaucoma care. *Eye,* **25**, pp. 1575-1580.

WESSELINK, C. and JANSONIUS, N.M., 2014. Predicting and Preventing Visual Impairment and Blindness by Incorporating Individual Progression Velocity in Glaucoma Care. *Investigative Ophthalmology & Visual Science*, **55**(7), pp. 4470-4474

WESTLAKE, W., 2000. Another look at visual standards and driving - Better tests are needed to determine driving ability. *British Medical Journal,* **321**, pp. 973-974.

WHITTAKER, S.G. and LOVIE-KITCHIN, J.E., 1993. Visual Requirements for Reading. *Optometry and Vision Science,* **70**(1), pp. 54-65.

WILD, J.M., DENGLER-HARLES, M., SEARLE, A.E.T., O'NEILL, E.C. and CREWS, S.J., 1989. The influence of the learning effect on automated perimetry in patients with suspected glaucoma. *Acta Ophthalmologica Scandinavica,* **67**(5), pp. 537-545.

WILD, J.M., HUSSEY, M.K., FLANAGAN, J.G. and TROPE, G.E., 1993. Pointwise Topographical and Longitudinal Modeling of the Visual Field in Glaucoma. *Investigative Ophthalmology & Visual Science,* **34**, pp. 1907-1916.

WILD, J.M., PACEY, I.E., HANCOCK, S.A. and CUNLIFFE, I.A., 1999. Between-Algorithm, Between Individual Differences in Normal Perimetric Sensitivity: Full Threshold, FASTPAC, and SITA. *Investigative Ophthalmology & Visual Science,* **40**, pp. 1152-1161.

WISHART, J., KONDO, T. and ELDERTON, E.M., 1931. The Mean and Second Moment Coefficient of the Multiple Correlation Coefficient, in Samples from a Normal Population. *Biometrika,* **22**(3/4), pp. 353-376.

WORLD HEALTH ORGANISATION, 2007. *Vision 2020 The Right to Sight - Global Initiative for the elimination of avoidable blindness - Action plan 2006-2011.* Geneva, Switzerland: WHO Press.

ZHENG, Y., WONG, T.Y., MITCHELL, P., FRIEDMAN, D.S., HE, M. and AUNG, T., 2010. Distribution of ocular perfusion pressure and its relationship with open-angle glaucoma: the Singapore Malay Eye Study. *Investigative Ophthalmology & Visual Science,* **51**, pp. 3399-3404.

ZHU, H., CRABB, D.P., SCHLOTTMANN, P.G., LEMIJ, H.G., REUS, N.J., HEALEY, P.R., MITCHELL, P., HO, T. and GARWAY-HEATH, D.F., 2010. Predicting Visual Function from the Measurements of Retinal Nerve Fiber Layer Structure. *Investigative Ophthalmology & Visual Science,* **51**, pp. 5657-5666.

ZHU, H., RUSSELL, R.A., SAUNDERS, L.J., CECCON, S., GARWAY-HEATH, D.F. and CRABB, D.P., 2014. Detecting Changes in Retinal Function: Analysis with Non-Stationary *Weibull* Error Regression and Spatial Enhancement (*ANSWERS*). *PLoS One,* **9**(1), pp. e85654.

List of Publications

Peer-reviewed publications

SAUNDERS, L.J., RUSSELL, R.A. and CRABB, D.P., 2012. Practical landmarks for visual field disability in glaucoma. *British Journal of Ophthalmology,* **96**(9), pp. 1185-1189.

SAUNDERS, L.J., RUSSELL, R.A. and CRABB, D.P., 2012. The Coefficient of Determination: What Determines a Useful R^2 Statistic? *Investigative Ophthalmology & Visual Science,* **53**(11), pp. 6830-6832.

SAUNDERS, L.J., RUSSELL, R.A. and CRABB, D.P., 2013. Detecting Visual Field Progression in Glaucoma – Using the Right Tools for the Job, *European Ophthalmic Review,***7**(1), pp. 20-26.

ZHU, H., RUSSELL, R.A., **SAUNDERS, L.J.**, CECCON, S., GARWAY-HEATH, D.F. and CRABB, D.P., 2014. Detecting Changes in Retinal Function: Analysis with Non-Stationary *Weibull* Error Regression and Spatial Enhancement (*ANSWERS*). *PLoS One,* **9**(1), pp. e85654.

SAUNDERS, L.J., RUSSELL, R.A., KIRWAN, J.F., MCNAUGHT, A.I. and CRABB, D.P., 2014. Examining Visual Field Loss in Patients in Glaucoma Clinics During Their Predicted Remaining Lifetime. *Investigative Ophthalmology & Visual Science,* **55**, pp. 102-109.

SAUNDERS, L.J., RUSSELL, R.A., KIRWAN, J.F., MCNAUGHT, A.I. and CRABB, D.P., 2014. Author Response: Predicting and Preventing Visual Field Impairment and Blindness by Incorporating Individual Progression Velocity in Glaucoma Care. *Investigative Ophthalmology & Visual Science,* **55**, pp. 4475.

SAUNDERS, L.J., RUSSELL, R.A. and CRABB, D.P., 2014. Detecting Visual Field Progression in Glaucoma – Using the Right Tools for the Job, *US Ophthalmic Review,***7**(1), pp. 32-38.

SAUNDERS, L.J., RUSSELL, R.A. and CRABB, D.P., 2015. Measurement Precision in a Series of Visual Fields Acquired by the Standard and Fast Versions of the Swedish Interactive Thresholding Algorithm: Analysis of Large-Scale Data From Clinics, *JAMA Ophthalmology,* **133**(1): 74-80.

SAUNDERS, L.J., ZHU, H., BUNCE, C., DORÉ, C.J., FREEMANTLE, N. and CRABB, D.P., 2015. Ophthalmic statistics note 5: diagnostic tests – sensitivity and specificity, *British Journal of Ophthalmology,* **Online First**.

BURTON, R., **SAUNDERS, L.J.** and CRABB, D.P., 2015. Areas of the visual field important during reading in Glaucoma, *Japanese Journal of Ophthalmology,* **Online First**.

Conference presentations (all read papers)

International Conferences

SAUNDERS, L.J., RUSSELL, R.A. and CRABB, D.P., 2014. 'Standard or Fast? - Differences in precision between SITA Standard and SITA Fast testing algorithms', paper presented at *The Association for Research in Vision and Ophthalmology Annual Meeting*, Orlando FL, May 2014

SAUNDERS, L.J., RUSSELL, R.A. and CRABB, D.P., 2013. 'Standard or Fast? – Comparing the relationship between variability and sensitivity in SITA Standard and SITA Fast using Large-Scale Longitudinal data', paper presented at *The North Atlantic Perimetry Society Meeting*, Chapel Hill NC, October 2013

National Conferences

SAUNDERS, L.J., RUSSELL, R.A. and CRABB, D.P., 2012. 'Standard or Fast? – Evaluation of the precision of SITA Standard and SITA Fast and their utility in detecting glaucomatous visual field progression', paper presented at *United Kingdom and Eire Glaucoma Society Meeting*, Bristol, November 2014

SAUNDERS, L.J., RUSSELL, R.A. and CRABB, D.P., 2014. 'Avoiding a blind alley – Examining vision loss in patients with glaucoma during their predicted remaining lifetime', paper presented at *Royal Statistical Society Young Statisticians Meeting*, London, July 2013

SAUNDERS, L.J., RUSSELL, R.A. and CRABB, D.P., 2012. 'Legally Unfit to Drive – how do clinical measurements of glaucoma damage correspond with this landmark?', paper presented at *United Kingdom and Eire Glaucoma Society Meeting*, Edinburgh December 2012

University Presentations

SAUNDERS, L.J., 2014. 'Need for speed – does using a faster test for monitoring glaucoma impact on time to detect disease progression?', paper presented at *City University 2nd Annual School of Health Sciences Doctoral Student Conference*, London, April 2014

SAUNDERS, L.J., 2014. 'No more turning a blind eye – visual fields and visual function in glaucoma', paper presented at *City University 3 minute thesis competition*, London, March 2014

SAUNDERS, L.J., 2013. 'Avoiding a blind alley – Examining vision loss in patients with glaucoma during their predicted remaining lifetime', paper presented at *City University 3rd Annual Research Symposium*, London, June 2013

SAUNDERS, L.J., 2013. 'Avoiding a blind alley – Examining vision loss in patients with glaucoma during their predicted remaining lifetime', paper presented at *City*

University 1st Annual School of Health Sciences Doctoral Student Conference, London, April 2013

SAUNDERS, L.J., 2012. 'Practical Landmarks for Visual Field Disability in Glaucoma', paper presented at *City University 2nd Annual Research Symposium*, London, June 2012