A COMPARISON BETWEEN TWO FAST STRATEGIES, SPARK AND SITA IN THRESHOLD PERIMETRY

SAY KIANG FOO

Doctor of Optometry

ASTON UNIVERSITY

September 2018

Aston University
A Comparison between Two Fast Strategies, SPARK and SITA in Threshold Perimetry
Say Kiang Foo
Doctor of Optometry
2018

This thesis investigated the extent of differences between a newer fast threshold strategy, SPARK and the “gold standard” strategy, Swedish Interactive Threshold Algorithm (SITA).

A between-visit comparative study in subjects with glaucoma and suspected glaucoma, subjects with cataract and healthy controls showed SPARK had a lower inter-test variability and was less influenced by the perimetric-experience of subjects compared to SITA. Pointwise analysis found reduction of superior peripheral sensitivity with SPARK in the second visit. Lesser unreliable visual field (VF) results were obtained with SPARK possibly due to its fixation loss monitoring methods and shorter test duration.

SPARK produced about 40% of time-saving and higher estimated sensitivity as compared to SITA in normal subjects. Higher estimated sensitivity may be attributed to reduced fatigue effect, instrumentation and applied algorithms. An acceptable agreement was achieved for mean sensitivity (MS) between the strategies and the bias decreased towards old age.

High bias and large limits of agreement of mean deviation (MD) were found between SPARK and SITA in glaucoma patients and the agreements of global indices between strategies were not achieved. Overestimation of MD and underestimation of pattern standard deviation (PSD) with SPARK led to underestimation of glaucoma severity level and poorer diagnostic sensitivity compared to SITA. The overestimation of MD was also found in cataract patients which led to the masking of diffuse loss. A depressed normative database was possibly used in SPARK algorithm. Between-strategy agreement was not found for the global indices in cataract patients.

Pointwise bias differences of threshold estimates between SPARK and SITA were found to have an almost similar progressive pattern with highest bias in nasal field decreasing towards the temporal field in either glaucoma, cataract or healthy subject which demonstrated a possible systemic difference between strategies. Hence, SPARK is yet to be an alternative strategy to SITA.

Fatigue effect, inter-test variability, global indices, mean deviation, diffuse loss

To

Chin Voon & Arinne

ACKNOWLEDGEMENT

I would like to thank my supervisor Dr Rebekka Heitmar for her guidance, encouragement and help. I would also like to thank my associate supervisor, Dr Robert Cubbidge for his help and guidance especially in the preparation of the thesis proposal and research ethics approval.

I would also like to extend my appreciation to the following individuals: Dr Khor Sim Ee and Associate Professor Dr Sushil Kumar Vasudevan for their help with the recruitment of glaucoma patients; Ms. Caroline Brocklebank for her help related to the official matters in School of Life and Health Science, Aston University; Prof Manuel Gonzalez de la Rosa and Dr Gagyi Palffy Zoltan for helping me to further understand about SPARK and Oculus perimetry; all my colleagues from Faculty of Optometry and Vision Sciences, SEGi University for their supports and helps especially in recruiting subjects and sharing my workloads to spare me for this research.

I am grateful to all the patients and subjects who participated in this study and also recommended friends and family for this research.

Finally, I would like to thank my lovely wife, Chin Voon and my daughter, Arinne for their tolerance, supports and understanding throughout the years.

LIST OF CONTENTS

LIST OF ABBREVATIONS

°	degree
%	Percent
>	more than
⩾	equal or more than
<	less than
⩽	equal or less than
/	or / per / divided by
AGIS	Advanced Glaucoma Intervention Study
asb	Apostilb
BC	Before Christ
BCVA	Best corrected visual acuity
cd/m^2	Candela per square metre
CIGTS	Collaborative Initial Glaucoma Treatment Study
cm	centimetre
CNTGS	Collaborative Normal-tension Glaucoma Study
CPSD	Corrected pattern standard deviation
D	Dioptre
dB	Decibel
DS	Dioptre sphere
DLS	Differential light sensitivity
DC	Dioptre cylinder
EMGT	Early Manifest Glaucoma Treatment
ERF	Error related factor
et al.	and others
FDT	Frequency-doubling technology perimetry
FL	Fixation loss

FN	False negative
FP	False positive
FT	Full Threshold
GATE	German Adaptive Thresholding Estimation
GHT	Glaucoma Hemifield test
Hb	Haemoglobin
HFA	Humphrey Field Analyzer
HPA	Hodapp-Parrish-Anderson
HRT	Heidelberg retinal tomography
IOP	Intraocular pressure
IPS	Imaging and Perimetry Society
LoA	Limits of agreement
LOCS III	Lens Opacities Classification System III
LV	Loss variance
MAP	Medmont automated perimeter
MD	Mean deviation
MDf	Mean defect
mm^2	square millimetre
ms	millisecond
MS	Mean sensitivity
NAPDP	Number of abnormal pattern deviation points
NC	Nuclear color
NO	Nuclear opalescence
NTG	Normal tension glaucoma
OCT	Optical coherence tomography
OHTS	Ocular Hypertension Treatment Study
ONH	Optic nerve head

PACG	Primary angle-closure glaucoma
PD	Pattern deviation
pg	page
POAG	Primary open-angle glaucoma
PSC	Posterior subcapsular cataract
PSD	Pattern standard deviation
RNFL	Retinal nerve fibre layer
s	second
SAP	Standard automated perimetry
SD	Standard deviation
SE	Standard error
SITA	Swedish Interactive Threshold Algorithm
SS	SITA Standard
SF	SITA Fast
SP	SPARK Precision
TD	Total deviation
TOP	Tendency-Oriented Perimetry
VA	Visual acuity
VF	Visual field
VFI	Visual field index
ZATA	Zippy Adaptive Threshold Testing
ZEST	Zippy Estimation by Sequential Testing

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

BACKGROUND ON FAST THRESHOLD PERIMETRY USING SITA AND SPARK

1.1 Perimetry

Perimetry is the procedure or technique used to measure the visual field (VF) of an eye. The VF is regarded as the area in which objects are visible when the fixation of eye is fixed in one direction with stationary head and body (IPS, 2012). Generally, VF is described to be the widest on the temporal side up to 90° and narrower on the nasal and superior due to the extent of the nose and the bone of the orbit. Prominent eyebrows, deep positioning of eyeballs or large nose may restrict the VF of an eye (Meyer et al.,1993) and the VF is larger when using both eyes. The VF within the 30° extension of the central fixation is called central field and the VF outside of the central field is called peripheral field (Anderson and Patella, 1999). A physiological blind spot is found in each eye which is located 15° temporally and 1.5° inferiorly from the central fixation. It has a size of about 7.5° in diameter and its centre is 1.5° inferiorly from the horizontal midline (Armaly, 1969). It corresponds to the location of the optic nerve head (ONH) of the eye which has no photoreceptors.

The first investigation of VF was reported by Hippocrates in approximately late fifth century BC which described hemianopsia (Lascaratos and Marketos, 1988; Johnson et al., 2011; Magnus, 1998). The shape of the VF was reported as circular by Ptolemy who attempted to quantify VF in 150 BC (Thompson and Wall, 2008; Lascaratos and Marketos, 1988). But the first exact measurement of VF was made by Thomas Young in 1801 who described the dimensions of the VF as 50° superiorly, 60° nasally, 70° inferiorly and 90° temporally (Thompson and Wall, 2008; Lascaratos and Marketos, 1988). Purkinje showed that the dimensions of VF were a bit larger in all directions when more detectable targets were used (Thompson and Wall, 2008; Lascaratos and Marketos, 1988). Visual field testing has become an essential examination in clinical ophthalmology started from the contribution of Albrecht von Graefe in 1856 when

quantitative VF examinations were obtained (Schiefer et al., 2005; Johnson et al., 2011). Tangent screen or Bjerrum screen is one of the earliest quantitative perimetric methods which is still used nowadays. It uses a uniform dark flat screen and small targets which are available in white or other colours to be presented on the tip of a dark wand to map out VF sensitivity. Due to this method only being able to measure central 30° radius of VF, the arc perimeter was developed to evaluate the full extent of the peripheral VF, but a uniform background could not be achieved for the full VF. It was not until the bowl perimeter was introduced which allowed stimuli to be projected onto a uniform background to determine the minimum increment of light needed to detect the target. Hans Goldmann was instrumental in developing the type of hemispherical bowl perimeter which also has a moving optical projection system (Johnson et al., 2011). Varieties of stimulus size and luminance levels are available in this Goldmann perimeter and the relationships among the size, luminance and VF locations were well-studied by Goldmann (Thompson and Wall, 2008; Lascaratos and Marketos, 1988). This perimeter has become the most common perimeter used for performing manual quantitative perimetry (Thompson and Wall, 2008; Aulhorn and Harms, 1972). The first fully automated perimeter known as Octopus (Haag-Streit, Koeniz, Switzerland) was introduced (Fankhauser et al., 1972; Fankhauser, 1982; Spahr and Fankhauser, 1974; Gloor, 2009) in the 1970s and it can be considered as the beginning of the modern perimetry which is not only using automated testing but also allows computerised test analysis. The development of Humphrey Field Analyzer (HFA) (Carl Zeiss Meditec, Dublin, CA) started in the 1980s and with the introduction of fast threshold strategy, Swedish Interactive Threshold Algorithms (SITA) (Bengtsson et al., 1997; Bengtsson et al., 1998; Bengtsson and Heijl, 1999), it has become the most commonly used automated perimeter in the recent decades.

1.2 Static perimetry

In static perimetry, stationary stimuli with constant size are presented at different location of the VF and the luminance of the stimuli is modified until threshold for the dimmest target the patient can see 50% of the time at each of the test locations is found (Sample et al., 2011;

Allingham et al., 2011). Static perimetry is one of the quantitative perimetries but manually operating static perimetry is very time-consuming. Static perimetry has become more widely used after the automated perimeter was introduced such as the Octopus Automated Perimeter (Fankhauser et al., 1972), Fieldmaster (Keltner et al., 1979) and Dicon (Mills, 1984) which allows standardized testing and shorter duration with minimal examiner bias.

There are factors that can affect the VF examination finding such as background illumination, stimulus duration, stimulus luminosity and the size of the stimulus (Nouri-Mahdavi, 2014; Khuu and Kalloniatis, 2015). The standard unit of measurement in the VF is the differential light sensitivity (DLS) which is the threshold of perception of a stimulus relative to its background illumination (Schiefer et al., 2005). Standard automated static perimetry measures DLS where the stimulus of varying luminance is presented on a specified background luminance (Sample et al., 2011) which is ranged from 1 to 10 cd/m^2.

1.2.1 Background Illumination

Most of the perimetric devices use 31.5 apostilbs (asb) or 10 cd/m^2 as the background luminance and it is the luminance level that Weber's law holds (Wall & Johnson, 2005). Weber's law states the ratio of the increment threshold to the background intensity is a constant (Schiefer et al., 2005; Aulhorn and Harms, 1972). There are a few advantages of using this background luminance level: a) it requires minimal pre-test adaptation time as it is close to offices and waiting room lighting conditions; b) it is comfortable to most of the patients; c) the least amount of response variability from patients at this background luminance; d) factors that affect the amount of light reaching the retinal such as pupil size, transmission level through ocular media, have equal effects on the stimulus and background luminance (Wall and Johnson, 2005; Spry and Harper, 2010; Johnson, 2013). A low background luminance and a large and high stimulus luminance can maximize the ability of stimulus detection. Longer dark adaptation time is required if very low background luminance is used which is not viable for routine clinical testing. Moreover, the unstable adaptation state is found for the whole VF which

is affected by red-cone density distributions, pupil sizes and spatial summation properties (Aulhorn and Harms, 1972; Greve, 1973; Johnson et al., 1981). High level of background luminance will cause discomfort, glare effects and false responses (Anderson et al., 2009; Fankhauser and Haeberlin, 1980). Older version of Octopus perimeter is one of the perimeters that uses 1 cd/m^2 as the background luminance which is under mesopic condition but it claimed to be more sensitive to mid-peripheral region field damage (Flanagan, 2009).

1.2.2 Stimulus Luminosity

Stimulus intensity in automated static perimetry is quantified by using a decibel (dB) scale which describes the difference of the magnitude of the luminance relative to a specified reference level in logarithmic scale. The maximum stimulus intensity capably produced by each instrument is used as its reference level of luminance. Zero dB is referred to the maximum luminance of the stimulus whereas 1 dB is equal to 1/10 log unit of attenuation of maximum available stimulus luminance (Reddy, 2010). There is no negative dB value as it is not possible to have stimulus luminance higher than maximal value. By using the logarithmic scale, a wide range of the luminance value (8 logarithmic units) can be used to quantify the range of the human visual system which can be as high as 1,000,000 asb. The maximum stimulus luminance for Goldmann perimeter is 1,000 asb, HFA (Carl Zeiss Meditec, Dublin, CA) is 10,000 asb, Oculus Centerfield and Twinfield (Oculus, Wetzlar, Germany) are 1,000 asb**,** Octopus 101 and 300 are 1,000 asb and 4,800 asb respectively (Weijland et al., 2004). Another Oculus perimeter (Easyfield) and the latest Octopus perimeter (Octopus 900) (Haag-Streit, Koeniz, Switzerland) are reported to be 10,000 asb. Due to the maximum and minimum intensities of stimulus light vary among static automated perimeters, the decibel scales between perimeters are not comparable (Pye et al., 1999).

1.2.3 Stimulus Size

The stimuli sizes available in both HFA and Oculus Twinfield are Goldmann size I, III and V (largest) but the standard stimulus size used in automated static perimetry is size III. It has a

size of 4 mm^2 (0.43°). Size V is larger with a size of 64 mm^2 (1.72°) usually used in cases with macular disorders and advanced glaucoma (Reddy, 2010). Recent study showed by using stimuli size V, the lower limit of the reliable stimulus range does not fall under 15 to 19 dB but higher sensitivity is found at the same location (Gardiner et al., 2015) which allows more reliable and less variability threshold estimation in the later stage of glaucoma (Wall et al., 1997; Wilensky et al., 1986). Larger stimuli may be less sensitive to the detection of early and small VF defects but Wall et al. (2013) had shown that the ability of stimuli size V to detect early field defects is as good as size III and poorer sensitivity in detecting mild glaucomatous loss was found when stimulus size VI (3.44°) was used. The stimuli size and luminance are the common variable parameters that can be modified to increase the visibility of the stimulus. A large and dim stimulus can be as visible as a smaller but brighter stimulus. This relationship between the stimulus size and its intensity can be described by spatial summation which is influenced by retinal location (Sloan, 1961; Sloan and Brown, 1962; Khuu and Kalloniatis, 2015a; Volbrecht et al., 2000; Vassilev et al., 2005; Wilson, 1970) and background adaptation (Barlow, 1958; Redmond et al., 2013; Lelkens and Zuidema, 1983).

Ricco's law states that the area (A) and intensity (I) of the stimulus are inversely proportional (Ricco, 1877) as long as the stimuli size is smaller than 10' (Goldmann stimuli size I and below). This size is small enough to be located within the receptive fields of a single retinal ganglion cell (Schiefer et al., 2005). Reduction of the stimulus intensity will need the stimulus area to be increased to achieve a similar sensation. Redmond et al. (2010) showed the largest stimulus size that Ricco's law holds true (Ricco's area) is enlarged in early glaucoma to maintain the constant threshold as a consequence of ganglion cell loss. The Ricco's area is the area with complete spatial summation and it is increased with eccentricity to maintain a constant number of retinal ganglion cells (Volbrecht et al., 2000; Vassilev et al., 2005; Vassilev et al., 2003). Age is not a factor that influences the Ricco's area (Redmond et al., 2010a). Due to the difference of Ricco's area in glaucoma eyes as compared to healthy eyes, Rountree et al. (2018) had further showed that using area-modulated stimuli could provide advantage in detecting early

glaucoma and glaucomatous progression which produces better disease signals, reduced response variability and higher signal to noise ratio compared to conventional fixed-stimuli size used in standard automated perimeter (Goldmann stimulus size III).

On the other hand, Piper's law is applied if the stimulus size is bigger than 10' (Goldmann stimulus size II and above) and it states that stimulus intensity is reciprocal to the square root of the stimulus area. Lateral inhibition of the adjacent neuron is induced if the stimulus size is larger than the size of the receptive fields of retinal ganglion cells (Schiefer et al., 2005).

1.2.4 Stimulus Duration

Temporal summation is described by Bloch's law which shows the reciprocal relationship between stimulus duration and threshold luminance for the duration which is shorter than critical duration (Bloch, 1885). The critical duration was reported in the range of 100 ms (Bruder and Kietzman, 1973; Saunders, 1975) and if the duration is reached, the contrast threshold is independent of duration. The limit for temporal summation is reached when the stimulus duration is more than 100 ms and more reliable and stable results will be produced (Aulhorn and Harms, 1972).

Stimulus duration of 200 ms is commonly used in the automated perimetry which is longer than the critical duration of complete temporal summation for the same type of stimulus but is shorter than the latency of saccadic eye movements which is approximately 250 ms (Mulholland et al., 2015a). Octopus perimeter is using 100 ms whereas Oculus and Humphrey are reported using 200 ms for their stimulus duration.

Mulholland et al. (2015) reported that the glaucoma eyes may have increased critical duration of complete temporal summation. Hence, the fundamental parameters used in the automated perimetry may need to be reviewed (Demirel, 2015).

1.3 Threshold Perimetry

Basically, threshold perimetry is found in static perimeter where the stimulus luminance is modified to determine the retinal threshold. In threshold perimetry, differential light sensitivity (DLS) is determined in individual stimulus locations across VF which represents the threshold of detecting stimulus 50% of the time it is presented (Schiefer et al., 2005; Mueller, 1951; Hallet, 1969; Weijland et al., 2004). The presented stationary stimuli have fixed size but vary in intensity. The reciprocal of the threshold is sensitivity (sensitivity = 1/threshold) (Wall and Johnson, 2005; Spry and Harper, 2010). The objective of the threshold perimetry is to quantify the VF. The threshold is expressed as sensitivity on a decibel scale, and the dimmer targets have higher decibel values. Zero decibel is set as the brightest stimulus that the perimeter itself can produce. The threshold in decibels is directly proportional to retinal sensitivity. It is measured in different locations of the retina in order to detect deviations from the values obtained from a population of normal eyes. Generally, in a normal eye, the fovea has the highest sensitivity and it drops rapidly between the fovea and 3°. Then it decreases gradually out to 30° and is followed with rapid dropping off beyond 50°. It is impractical to measure the threshold at a specific retinal location by testing multiple times with different stimulus intensities especially if the initial stimulus luminance is far from the threshold. Therefore, various threshold strategies such as staircase algorithms and technique using Bayesian predictive probability are used in static automated perimeters to estimate threshold.

As mentioned by Gonzalez de la Rosa and Gonzalez-Hernandez (2013), threshold is a probabilistic concept and its most representative value can be determined by averaging the threshold values obtained for a particular point. The first generation of algorithms used to determine the threshold estimates is by simply modifying the stimulus intensities in ascending or descending order until a threshold was crossed at all stimulus locations (Chaplin et al., 1973; Heijl and Krakau, 1975; Heijl, 1977). The full threshold strategy considered as the clinical gold standard technique in static threshold perimetry which is used in most glaucoma-related clinical trials. Octopus perimeter uses 4-2-1 Staircase or Bracketing method to determine the

retinal sensitivity and it is considered as the first generation of the algorithm (Bengtsson et al., 1997). In this algorithm, stimuli with 4-dB steps are presented until the patient gives a reversal response. It is followed by using 2-dB steps to get another reversal response and lastly using 1-dB step to obtain the threshold for that particular location.

Full threshold strategy in Humphrey Field Analyzer (HFA) applies 4-2 staircase strategy with the initial crossing in 4-dB step size and second crossing in 2-dB step size (Bebie et al., 1976; Johnson et al., 1993). The last–seen stimulus luminance is regarded as the threshold for that particular point. Full threshold strategy starts with determining the threshold in the four primary points that are located near the centre of each quadrant. By using the DLS of the four points, the initial stimulus of the immediate adjacent test points can be determined by referring to their local slope of the normal hill of vision in that area and threshold value at the primary points. This could save time rather than starting at the age-corrected normal level (Weijland et al., 2004; Wild et al., 1999).

The measured thresholds will be compared to threshold values from a database of age-matched patients at a particular stimulus location in the VF. Statistical analysis can be carried out to facilitate the diagnosis and monitoring of the VF. The examined area of VF in static automated perimetry usually is confined to the central 30°. Areas such as peripheral field between 30° and 60° and macula are also used in threshold perimetry.

1.4 Fast Threshold Strategy

The full threshold VF test usually takes longer time to complete the tests which could take up to about 15 minutes or more per eye with a 30-2 test point pattern even for a normal subject (Schimiti et al., 2002; Wild et al., 1999; Bengtsson et al., 1998; Budenz et al., 2002). It was also reported that Full Threshold static perimetry (HFA 30-2) requires 25% longer testing time than manual Goldmann kinetic perimetry among glaucoma patients (Trope and Britton, 1987). Theoretically, the more time is taken for multiple verifications and repetitive measurements in

the test, the accuracy of the threshold estimation will be higher (Heijl, 1977; Johnson and Nelson-Quigg, 1993). The severity of the VF defects could prolong the test duration (Wild et al., 1999; Roggen et al., 2001). However, longer testing time requires a certain level of patient's attention span and endurance. It can increase intratest and intertest variability (Bengtsson et al., 1998; Aoki et al., 2007; Wild et al., 1999), poor patient acceptance (Glen et al., 2014) and induce "fatigue effects" which inevitably compromise the accuracy of the threshold estimates (Johnson et al., 1988; Heijl and Drance, 1983; Gonzalez de la Rosa and Pareja, 1997; Hudson et al., 1994). Gonzalez de la Rosa and Pareja (1997) showed threshold reduction of 0.08-0.1 dB of sensitivity with each extended minute of testing time. Lower threshold sensitivities were found in longer duration perimetric test using staircase algorithm in normal subjects (Searle et al., 1991; Heijl, 1977a; Hudson et al., 1994; Johnson et al., 1988; Langerhorst et al., 1987; Suzumura, 1988; Wildberger and Robert, 1988; Bengtsson et al., 1998) and patients with ocular hypertension (Langerhorst et al., 1987; Suzumura, 1988), glaucoma (Heijl, 1977a; Johnson et al., 1988; Langerhorst et al., 1987; Suzumura, 1988; Holmin and Krakau, 1979; Heijl and Drance, 1983) and optic neuropathy (Wildberger and Robert, 1988).

Methods of using the test results at the adjacent points to determine the initial stimulus brightness and adjusting the pace of the test by measuring patient reaction times only shorten up to 20% of the testing time (Heijl and Patella, 2002). Many fast threshold strategies are introduced in automated static perimeters to further improve the test brevity with minimal compromise on the accuracy of the threshold estimates and also to minimize the examination burden on patients. Commercially available fast threshold strategies are FASTPAC, SITA Standard and SITA Fast incorporated in HFA; Dynamic strategy, tendency-oriented perimetry (TOP), German Adaptive Thresholding Estimation (GATE) (Schiefer et al., 2009) in Octopus perimeters (Haag-Streit, Koeniz, Switzerland); Zippy Adaptive Threshold Testing (ZATA) and ZATA Fast in Henson perimeters (Elektron Technology, Santa Fe Springs, CA, USA); Zippy Estimation by Sequential Testing (ZEST) in Medmont perimeters (Medmont Pty. Ltd.,

Nunawading, Victoria, Australia); and SPARK (Gonzalez de la Rosa and Gonzalez-Hernandez, 2013) in Oculus perimeters (Oculus, Wetzlar, Germany).

The testing duration initially was shortened by increasing the step size and/or reduce the number of threshold crossing. FASTPAC is the first fast threshold strategy available in HFA which accelerates the testing procedure by using 3dB-steps throughout the test without reversal response (Flanagan et al., 1993; Flanagan et al., 1993a). The dynamic strategy in Octopus also modifies the step size but it is according to the subject's sensitivity at that location which uses larger steps in areas with low sensitivity and smaller steps in areas with high sensitivity (Weber, 1990; Anderson and Johnson, 2006; Weber and Klimaschka, 1995). Both strategies still require to use a high number of stimulus hence the testing duration is still relatively time-consuming. FASTPAC was found to be less accurate (O'Brien et al., 1994; Glass et al., 1995) and had larger intra-test variability (Flanagan et al., 1993; Schaumberger et al., 1995) while the dynamic strategy showed lower reproducibility and higher variance in regions with relative defects (Vivell et al., 1991; Weber and Klimaschka, 1995) when compared to the Full Threshold algorithm.

With the improvement of the software that helps in the acquisition of the threshold estimate, application of complex statistical methods and advanced development in psychophysical studies, strategies using Bayesian inference and maximum likelihood principles such as SITA and ZEST (King-Smith et al.,1994; Turpin et al.,2002; Turpin et al., 2003; Vingrys and Pianta, 1999) were introduced. Both are forecasting techniques that improve the accuracy and efficiency of the VF threshold estimations (Johnson, 2013).

There are also strategies using the relationship between the threshold values of neighbouring VF test locations such as TOP (Lachkar et al., 1998; Maeda et al., 2000; Morales et al., 2000) to estimate the threshold values. But TOP was found less effective in detecting small VF losses

and shallower scotomas (Anderson, 2003; King et al., 2002) despite requires only half of the testing time by SITA Fast (Wadood et al., 2002).

More recent strategy, German Adaptive Thresholding Estimation (GATE) which does not use the prior knowledge from normal and glaucomatous populations but additional information from the previous examinations is used for threshold estimation. It uses modified 4-2dB to determine DLS. It produced shorter testing time than SITA Standard with comparable accuracy and test-retest reliability (Schiefer et al., 2009). ZATA also uses prior data and Bayesian inference likes SITA but with more flexible terminating criteria (Harvey, 2011).

All of these strategies have been developed to provide more reliable and stable results with minimal intra-test and inter-test variability which helps in the diagnosis and monitoring VF related ocular diseases especially glaucoma.

1.4.1 Swedish Interactive Threshold Algorithms (SITA)

The Swedish Interactive Threshold Algorithm (SITA) was developed in the mid-1990s by a research group in Sweden and incorporated in Humphrey Field Analyzer (HFA) (Carl Zeiss Meditec, Dublin, CA) (Bengtsson et al., 1997; Bengtsson et al., 1998; Bengtsson and Heijl, 1999). Since then, it was used in numerous clinical trials and is regarded as the clinical gold standard for VF examination. SITA was proven to provide faster VF testing time without compromising the accuracy of the result when compared to the Full Threshold and FASTPAC algorithms they replaced (Bengtsson et al., 1997; Bengtsson et al., 1998; Bengtsson and Heijl, 1998; Bengtsson and Heijl, 1998a; Sharma et al., 2000; Sekhar et al., 2000). Because of the accuracy and reproducibility of the results, it has been used in many clinical studies to compare new fast threshold strategies or devices (Landers et al., 2010; King et al., 2002; Schieffer et al., 2009; Capris et al., 2008).

SITA can be used for central 30-2, 24-2, 10-2 or peripheral 60-4 test points. It starts with using 4-2 staircase strategy to estimate threshold values for the four primary points with each located at 12.7° from the fixation point in each quadrant of the field. The threshold values of these points will be used to determine the initial stimulus luminance for their adjacent points in the VF. It is based on the prior knowledge of normal and glaucomatous VFs including age-corrected normal threshold values at each test point and random inter-subject VF variability to estimate threshold values and the measurement errors of the threshold values using Bayesian probability in VF models (Bengtsson et al., 1997; Cubbidge, 2005; Schieffer et al., 2009). The correlation between threshold values at adjacent locations is also taken into consideration. During the test, SITA continuously updates and modifies the Bayesian posterior probability distribution at each point based on patient's responses (Bengtsson et al., 1997; Bengtsson and Heijl, 1998; Bengtsson et al., 1998; Artes et al.,2002). The peak of the distribution represents the threshold value at each point that has the largest posterior probability whereas the width of the distribution describes the accuracy of threshold estimates. The testing is stopped at either the third reversal response or the measurement errors have been reduced to a predetermined level specified by error-related factor (ERF) (Olsson and Rootzen, 1994) whichever comes first. Error-related factor can only be used to terminate the testing when there is at least one crossing of the threshold and it cannot be changed (Bengtsson et al., 1997; Bengtsson and Heijl, 1998; Delgado et al., 2002). The threshold value at each testing location is recalculated at the end of the examination using all the information obtained during the examination. The details of the post-examination process are proprietary and unpublished (Schiefer et al., 2009).

Apart from that, the number of stimulus presentation is further reduced by eliminating catch trial for false-positive rates and instead of using the history of response-times within the test to determine the rates (Olsson et al., 1997). The false negative rate is also determined in SITA but decreased numbers of catch trials are used. SITA also adapts the inter-stimulus interval to the patient's reaction time. SITA only repeats the test for a test point when its estimated

threshold value is more than 12 dB from the initial estimate of threshold whereas the Full Threshold requires only more than 4 dB to repeat the test (Bengtsson et al., 1997; Olsson et al., 1997).

SITA has two testing strategies i.e. SITA Standard and SITA Fast. In SITA Fast, the testing can only be terminated by ERF if there is at least one positive response is recorded with the condition of the measurement error is smaller than the predetermined accuracy level specified by ERF which means crossing the threshold is not necessary for the termination of the testing. A single reversal staircase with 4 dB step size is used when the measurement errors are larger and it is different compared to SITA Standard which uses full staircase 4-2 dB with two reversals. The differences between these two strategies which include the predetermined level of accuracy for the termination of the test enable SITA Fast to produce shorter testing time (Bengtsson and Heijl, 1998a).

The objective of SITA is to reduce the duration of the VF test without affecting the quality of the test result compared to its predecessor algorithms Full Threshold (FT) and FASTPAC. Numerous comparisons of SITA Standard and FT have been reported since it was introduced by Bengtsson et al. (1997). A computerized simulation was used in the beginning to optimize the differences between the two strategies. In this simulation, Bengtsson et al. (1997) reported that the number of stimulus exposures was reduced by an average of 29% in normal fields and 26% in glaucomatous fields compared to FT. Bengtsson et al. (1998) started the actual clinical comparison of SITA with FT and FASTPAC in normal subjects using 30-2 test pattern. The clinical test time using SITA was reduced by 50% and 16% compared to FT and FASTPAC respectively with comparable test-retest variability to FT but significantly better than FASTPAC (Bengtsson et al., 1998). Reduced testing time could minimize the influence of the fatigue effect in SITA which produced higher threshold values in normal subjects compared to FT and FASTPAC (Bengtsson et al., 1998). Comparison in glaucoma subjects was also conducted by Bengtsson and Heijl (1998) with approximately the same percentage of test time reduction

found using SITA (46% for SITA Standard and 15% for FASTPAC) compared to FT but no significant difference was found for the test-retest variability among the three strategies. Higher threshold values were obtained with SITA and FASTSPAC compared to FT and the differences were larger than the normal subjects. This could be due to the visual fatigue is more prominent in glaucoma eyes than normal eyes. The differences mainly were detected on the test points with threshold values of >20 dB. These pieces of evidence have shown that SITA could be used to replace FT in glaucoma management with shorter testing time.

Another family member in SITA test strategies is SITA Fast which was developed to provide even shorter testing time with equal quality as FASTPAC (Bengtsson and Heijl, 1998a). Its stimulus sequences are terminated earlier than SITA Standard by using higher ERF cut off value and without the necessity of threshold crossing. SITA Fast reduced the clinical test time in patients with manifest and suspect glaucoma by 66% and 47% compared with Full Threshold and FASTPAC respectively (Bengtsson and Heijl, 1998a). It has low test-retest variability and defects found were often deep and more localised compared to those results by FT and FASTSPAC. Shallow field defects tend to be less prominent or may be missed by SITA Fast compared to FT and FASTPAC (Bengtsson and Heijl, 1998a).

Numerous published reports were followed suit after the introduction of the SITA strategies (Bengtsson et al., 1997; Bengtsson et al., 1998; Bengtsson and Heijl, 1998; Bengtsson & Heijl, 1998a). Firstly, looking at the test duration, significant reduction by 49 to 56% were reported when using SS compared to FT in normal subjects (Tsuji et al., 1998; Nordmann et al., 1998; Wild et al., 1999; Shirato et al., 1999; Remky and Arend, 2000; Budenz et al., 2002; Wall et al., 2001) while relatively lower range (45 to 50%) were reported in glaucoma patients (Inazumi et al., 1998; Nordmann et al., 1998; Wild et al., 1999a; Shirato et al., 1999; Sharma et al., 2000; Budenz et al., 2002) except Remky and Arend (2000) showed only 40% of test time reduction whereas another comparison by Sekhar et al. (2000) showed it can be up to 53% of reduction. There was no test time reduction in advanced glaucoma patients (Remky and

Arend, 2000) but a higher percentage of time saved was found in eyes with higher mean sensitivity or required longer testing time (Shirato et al., 1999). SITA Standard (SS) also showed time-saving in patients with optic neuropathies or hemianopia which 46% and 42% of the testing time were saved respectively when compared to FT (Wall et al., 2001).

Test duration was reduced by 14.7% and 24% when SS was compared with FASTPAC in normal subjects according to Wild et al. (1999) and Roggen et al. (2001) respectively. Whereas in glaucoma patients, SS was able to reduce the test time by 16% and 17.7% reported by Roggen et al. (2001) and Wild et al. (1999a). Those results are close to the figures reported earlier by Bengtsson and Heijl (1998).

SITA Fast was first introduced with the comparison conducted only in glaucoma patients and the test duration was found to be much shorter than SS, with test time reduction by 66% compared to FT (Bengtsson and Heijl, 1998a). Some studies have compared SF and FT in normal subjects and have reported relatively higher percentages of test time reduction (70 to 73%) (Tsuji et al., 1998; Nordmann et al., 1998; Wild et al., 1999; Budenz et al., 2002). Comparison of SF and FT in glaucoma patients was also conducted in some studies with the highest reduction in test time found at approximately 71% and the lowest about 65% of test time when using SF (Inazumi et al., 1998; Nordmann et al., 1998; Wild et al., 1999a; Sekhar et al., 2000; Budenz et al., 2002). Even though SF was designed to have comparable accuracy to FASTPAC, it produced shorter test duration than FASTPAC by at least 50% in normal subjects (Wild et al., 1999; Roggen et al., 2001) and 48% in glaucoma patients (Wild et al., 1999a; Roggen et al., 2001). These findings are consistent with the earlier report by Bengtsson and Heijl (1998a). Moreover, SF can also be used in healthy children which its testing duration was shortened by 33% when compared to FASTPAC (Akar et al., 2008). Conway and colleagues (2014) had shown that SF could be recommended instead of SS and Full threshold on epileptic patients with exposure of vigabatrin therapy. Visual field loss can be accurately

mapped out in these patients even though SITA algorithms are based on the prior knowledge of normal and glaucomatous VF behaviours to estimate threshold values.

Both SITA strategies were used to compare with each other and test duration was shortened by 41% in normal subjects and 39.9% in glaucoma patients using SF compared to SS (Wild et al., 1999; Wild et al., 1999a). Pierre-Filho et al. (2006) showed a test time reduction of 36.5% when using SF compared to SS in the combination of normal subjects and glaucoma patients. The advantage of SF in terms of time-saving is more prominent when older normal subjects are examined compared to SS (Wild et al., 1999). A different result was found in glaucoma patients as age is not a function of test duration using SITA strategies but the severity of field loss is (Wild et al., 1999a; Roggen et al., 2001). The proportionate time increment was higher when using SITA strategies especially SF with a higher level of the severity field loss (Wild et al., 1999a).

Lower test-retest variability in SITA strategies was found in earlier studies (Bengtsson et al., 1998; Bengtsson and Heijl, 1998a). A similar outcome was also shown by Shirato et al. (1999) and no significant difference was reported between SS and FT. However, Artes et al. (2002) showed the test-retest variability using SS was lower compared with using FT for all the test points but SF produced higher test-retest variability for test points with sensitivity below 20 dB. Therefore, SF was regarded as a strategy that is not recommended for monitoring established VF loss in spite of the significant reduction of its test duration (Artes et al., 2002). Sekhar et al. (2000) also showed variable repeatability by using SF whereas SS and FT produced excellent repeatability by comparing intraclass correlation.

Higher threshold values found using SITA strategies were hypothesized to be due to reduced fatigue with shorter test time (Bengtsson and Heijl, 1998, Bengtsson and Heijl, 1999a) but this explanation is disputed by Artes et al. (2002) as higher threshold values also were found from computer simulations of SF strategy and also among primary seed points and their closest

neighbours which were examined early in the test. There is a claim of the higher sensitivity values in SITA which are not just due to fatigue but also might be related to the differences in the initial values for the staircase procedure as the initial values are usually higher than the true threshold (Shirato et al., 1999). There was a previous report also mentioned that it might be due to the methodological differences in threshold determination between the algorithms (Hirasawa and Shoji, 2016) i.e. FT uses the last-seen stimulus as threshold value (Flanagan et al., 1993) whereas SITA determines threshold value as a stimulus with 50% probability of seeing which also involves a postprocessing algorithm to finalise the threshold value (Bengtsson et al., 1997; Bengtsson et al., 1998). The differences between the threshold values among strategies were larger in testing points that have lower sensitivity in glaucoma patients (Sharma et al., 2000; Budenz et al., 2002; Budenz et al., 2002a; Aoki et al., 2007) but Shirato et al. (1999) showed the difference at each test point was not dependent on the test point position or the sensitivity of that point. The difference was also found to be independent of age (Wild et al., 1999). Mean sensitivity of SS was averagely 1 dB and 0.7 dB higher than that of FT and FASTPAC respectively but 0.9 dB lower than that of SF in glaucoma patients (Wild et al., 1999a). It was also 0.8 dB higher than that of FT but 0.5 dB lower than that of SF among normal subjects (Wild et al., 1999). The differences in the threshold value between algorithms were clinical insignificant even though there were statistically significant (Wild et al., 1999a; Budenz et al., 2002a). The strategy with the shortest test time produced the highest threshold values (Wild et al., 1999; Bengtsson and Heijl, 1999).

Comparison of global indices among the strategies showed higher MD using SS and SF compared to FT either in normal subjects or glaucoma patients (Budenz et al., 2002; Budenz et al., 2002a; Sekhar et al., 2000; Hirasawa and Shoji, 2016; Akar et al., 2008; Heijl et al., 2000) but there are also reports showed no significant difference (Roggen et al., 2001; Bengtsson and Heijl, 1999). There was no difference found in PSD (Budenz et al., 2002; Bengtsson and Heijl, 1999; Heijl et al., 2000) when SITA strategies were compared to FT or FASTPAC in glaucoma patients but SITA showed lower PSD in normal subjects (Roggen et

al., 2001; Akar et al., 2008; Budenz et al., 2002). The numbers of defects with pattern deviation probability maps were greater in SITA strategies compared to FT (Sekhar et al., 2000). With the higher threshold values, SS and SF could underestimate the VF defects and give artificial improvement if compared to other strategies (Nordmann et al., 1998).

The depth of the field defects was shallower using SS and SF compared to FT but the defect size found by SS was larger than FT (Budenz et al., 2002a; Aoki et al., 2007; Hirasawa and Shoji, 2016), whereas there was no difference in size found between SF and FT (Budenz et al., 2002a). The shallower depth and smaller defect field using SS were also reported by Sharma et al. (2000) but the changes were not statistically significant. Contradicting results reported by Wild et al. (1999a) showed that statistically deeper defect was found by SS compared to FT and it was more apparent when the defect depth increased. This possibly was due to the higher age-corrected confidence limits for normality in each test location as shown by Wild et al. (1999). Total and pattern deviation probability plots showed a greater number of defects found using SS and SF (Wild et al., 1999a) compared to FT which could be due to tighter normal confidence limits as a consequence of reduced fatigue in shorter test duration or the possibility of higher accuracy in SITA strategies (Bengtsson and Heijl, 1999). Shallower field defects were also shown in a study by Remky and Arend (2000) that relative scotomas detected with FT were shown normal reading with SS. It also explained why the field defects detected by SITA were mostly absolute defects (Remky and Arend, 2000). As such, glaucomatous VF defects detected by SITA will appear to be improved if FT was used earlier (Wild et al., 1999).

Inter-subject variability was also compared among SITA strategies, FT and FASTPAC and SITA strategies were shown to produce lower inter-subject variability. It could cause narrower normal limits in normal subjects and the possibility of earlier detection of VF defects statistically is higher than FT and FASTPAC (Wild et al., 1999). Bengtsson and Heijl (1999a) also showed SF and SS produced more than 40% and 30% respectively lower mean inter-subject threshold

variances compared to Full Threshold among normal subjects. Lower inter-subject variability using SS also found in suspected and glaucoma patients compared to FT (Aoki et al., 2007).

SITA strategies have been widely used in glaucoma diagnosis and their sensitivities and specificities were reported but the criteria used were variable. King et al. (2002) showed that sensitivities of SF in detecting VF loss in glaucoma patients ranged from 86.4 to 89.2% and specificities from 80.8 to 93.8%, based on criteria recommended by Hodapp-Parrish-Anderson (HPA). Wadood et al. (2002) also used the same diagnostic criteria but higher sensitivities (94.2 to 98.5%) and lower specificities (67.8 to 82.1%) of SITA Fast were found. These differences may be due to different groups of glaucoma patients who could have different severity of VF defects. The severity of glaucoma, location, eccentricity, age and mean pupil size could explain the variability of the VF test (Blumenthal et al., 2003). It was shown by Budenz et al. (2002) that the severity of the VF loss affects the sensitivity of the VF test using SITA strategies. Sensitivities of SS and SF on mild VF defect groups were found lower compared to severe VF defect groups.

Diagnosis of glaucoma also cannot just rely on one VF test. Repeated VF tests are required in order to confirm the diagnosis and to achieve higher specificity (Johnson et al., 2002; Gardiner and Crabb, 2002). Specificity was found lower in patients with no experience in VF test whose specificity with SS was reported as 38.1% compared to 63.2% with FT. Specificities were improved after the first test and no significant difference between SITA and FT was found (Schimiti et al., 2002). It was explained that FT did not produce the drastic change in specificity compared to SS could be due to the longer testing time using FT in the first test has provided more learning in preparation for the second test. It contradicts the direct comparison on the learning effect of FT and SS reported by Yenice and Temel (2005) in which SS was shown to have reduced learning effect compared to FT but specificity was not compared in the report. The learning effect of SS has been found to be correlated with age and education level (Aydin et al., 2015). Subjects who are more than 50 years old with an education level of below high

school showed a greater learning effect but Castro et al. (2008) found otherwise with no factors related to these changes.

High sensitivities of 92.68% for SF and 95.12% for SS were also reported by Sekhar et al. (2000) in glaucoma patients but the classification of their glaucoma patients was fully dependent on the results of FT strategy as their “gold standard”. Besides that, only the glaucoma hemifield test (GHT) was used to determine the VF loss in this study.

Liu et al. (2011) reported lower sensitivity for SS of 68.4% and even lower sensitivity for early glaucoma group of 46.4%. This was due to the different definition of glaucomatous eyes used. This study referred to the optical coherence tomography (OCT) retinal nerve fibre layer (RNFL) deviation map to classify glaucoma patients whereas studies mentioned earlier depend on the stereoscopic examination of a glaucoma specialist (King et al., 2002; Wadood et al., 2002). The lower sensitivities value could also be due to some of the early stage of glaucomatous eyes which already have changes in RNFL but yet to show any VF defects.

As SITA has shown good reliability and comparable accuracy to FT, it is widely used as the standard in threshold perimetry (Sharma et al., 2008; Jampel et al., 2011) and numerous newly introduced threshold strategies have used SS or SF as the comparison standard. Threshold strategies such as tendency-oriented perimetry (TOP) (King et al., 2002; Wadood et al., 2002; Rowe et al., 2014), the German Adaptive Thresholding Estimation (GATE) (Schiefer et al., 2009), and threshold perimetry in Medmont automated perimeter (MAP) (Medmont, Camberwell, Australia) (Landers et al., 2003; Landers et al., 2010) have been compared with SS or SF.

Tendency-oriented perimetry (TOP) produces much shorter testing time but it tends to underestimate the focal VF loss compared to SF (King et al., 2002) and lack of correlation shown in moderate to severe field defects between Octopus perimeter (dynamic strategy and

TOP) and Humphrey perimeter (SS and SF) (Rowe et al., 2014). German Adaptive Thresholding Estimation (GATE) produces similar threshold estimates compared to SS but it has the advantage of consistent examination duration it does not depend on the severity of the field defects whereas SS does (Schiefer et al., 2009). The threshold test using MAP correlates well with HFA SITA despite using larger step size (6 dB) and green stimulus light in dimmer background illumination for 100 test points (Landers et al., 2003; Landers et al., 2007). Threshold sensitivities estimated by MAP are approximately 5 dB lower than those by HFA SITA and their sensitivities are related linearly (Landers et al., 2010).

SITA has been used as a benchmark for diagnostic ability on glaucomatous field defect. More recent types of perimetry such as short-wavelength automated perimetry (SWAP) (Bengtsson and Heijl, 2006; Soliman et al., 2002; Girkin et al., 2000), Heidelberg contour perimetry (Fabrikantov et al., 2015), frequency-doubling technology (FDT) perimetry (Burgansky-Eliash et al., 2007; Liu et al., 2011; Kocabeyoglu et al., 2013; Lamparter et al., 2013; Leeprechanon et al., 2007; Racette et al., 2008; Sakata et al., 2007), flicker frequency-defined (FDF) perimetry (Calvo Perez et al., 2010), and Pulsar perimetry (Zeppieri et al., 2010) have been compared or determined their relationship with SITA. SITA is also incorporated into SWAP to optimize the diagnostic ability by shortening the testing time. SITA SWAP produced comparable diagnostic sensitivity with SF in patients with early glaucoma within a much shorter time compared to Full Threshold SWAP (Bengtsson and Heijl, 2006).

SITA Standard (SS) is regarded as the standard in the evaluation of glaucomatous functional damage despite some of the eyes may show structural changes at the level of the ONH or RNFL without evidence of VF loss (Gordon et al., 2002). A combination of examination of structural damage to the optic nerve and evaluation of visual function is commonly used for the diagnosis of glaucoma (Sharma et al., 2008). Studies on techniques used in optic nerve structural damage examination especially objective methods were using SS either to classify the study participants or/and compare the diagnostic accuracy individually or determine the

correlation between functional and structural examinations (Bozkurt et al., 2008, Lopez-Pena et al., 2011; Benitez-del-Castillo et al., 2008; Le et al., 2015; Bizios et al., 2011; Leite et al., 2012; Hirasawa et al., 2016; Fabrikantov et al., 2015).

SITA Standard was also used in children and the youngest subject reported by Donahue and Porter (2001) was 7 years old. Besides obtained 50% shorter testing time compared to FT, SS also demonstrated lower variability and comparable results including consistent reliability indices. However, this study did not test on all the same subjects for both SS and FT. Moreover, the possibility of a lower sample size of normal value database for young children in Humphrey software may produce some artifactual VF defects (Donahue and Porter, 2001). Akar et al. (2008) used SF on paediatric subjects for VF tests has shown promising results especially for children older than 8 years old who achieved significant high-reliability scores.

Another new family member of SITA strategies was recently introduced, SITA Faster which is intended to replace SF. SITA Faster produced 30.4% shorter testing time but identical results compared to SF (Heijl et al., 2019). Some modifications in SITA Faster are made to produce even faster VF test such as different starting stimulus intensities, one reversal for primary test points, SF used as prior distribution models, eliminating retesting procedure on non-seeing test points. false negative catch trial and blind spot locating step at the beginning of the test and the change of stimulus timing. The shorter procedure is more ideal for patients who may require more frequent VF tests in order to have better detection and monitoring of glaucoma.

1.4.2 SPARK

SPARK, denoting the flashes of the stimulus light and short duration of examination (Gonzalez de la Rosa and Gonzalez-Hernandez, 2013) is a new fast threshold strategy which uses the relationship between the distant test points of the VF which are connected by the tract of ganglion cell axons and the global optic nerve damage to speed up threshold examinations. It uses the probabilistic Bayesian dependence relations between the deviations in the

morphologic sectors described by Garway-Heath et al. (2000) (Figure 1.1) and functional sectors defined by Gonzalez de la Rosa et al. (2002) (Figure 1.2). The sequence of stimuli and the intensity of the next one in the test are determined by the most probable deviation.

Figure 1.1: The mapping of the visual field (A) with sectors in optic disc head (B) according to the study conducted by Garway-Heath et al. (2000).
Source: Garway-Heath et al. (2000)

Figure 1.2: Functional map of visual field defined by Gonzalez de la Rosa et al. (2002)
Source: Gonzalez de la Rosa et al. (2002)

A sample consists of 90,335 examinations using G1 program and TOP strategy of the Octopus 1-2-3 perimeter (Haag-Streit AG, Bern, Switzerland) was used to design SPARK. The sample is not limited to normal and glaucomatous eyes but it also includes 5,774 eyes with neurological, retinal or mixed diseases. It was intended to be used in the non-selected population (Gonzalez de la Rosa and Gonzalez-Hernandez, 2013). From the sample, six points with each point in relation to each sector shown in Figure 1.1 and 1.2 and their threshold deviations best correlated with the mean deviation were selected. The frequency of deviations at each point was recorded from the sample and a sequence of stimuli was designed to allow fast probabilistic estimation of the VF. Six stimulus points that are highly representative of entire VF are used in the first phase which usually lasts about 40 seconds. After that, the magnitude of the deviations from the normal age-corrected threshold values for all the rest of the points that were not examined can be estimated from the results of the six points using multiple regression equations. This first phase of the measurement serves as a starting point for three following phases of the test. The new deviation estimated is the result of the previous one plus or minus the standard error and it continues to do so for the deviation estimation of all the points in the sector. It is repeated in the following phases to further refine the local deviations with assigned sensitivities are continuously corrected according to the patient's responses. Finally, point-by-point analysis of the results obtained in each of the four phases is performed. The final result is obtained by using the median or the mean of the three most central values while the most extreme threshold estimate is rejected (Gonzalez de la Rosa and Gonzalez-Hernandez, 2013; Gonzalez de la Rosa, 2014). The first phase can also be considered as a short training phase for the inexperienced patient to achieve a stable result (Gonzalez de la Rosa and Gonzalez-Hernandez, 2013). There are a total of 75 to 80 stimuli presentations in all four phases which take approximately three minutes for the whole examination.

The SPARK strategy is available in Oculus Easyfield, Smartfield, Centerfield 2 and Twinfield 2 perimeters (Oculus Optikgeräte GmbH, Wetzlar, Germany) and it is designed not only to be

used in glaucomatous population but also in patients with neurological, retinal or mixed diseases. SPARK has a few types of examination according to the phases involved such as SPARK Precision, SPARK Quick and SPARK Training. SPARK Precision is a complete examination which involves all four phases while SPARK Quick only takes half of the time and only involves the first two phases. Basically, SPARK training is the first phase and it only lasts about 40 seconds. Test strategy SPARK-N is also available and used for neurological patients.

Gonzalez de la Rosa and Gonzalez-Hernandez (2013) designed this strategy based on a retrospective analysis using the results of over 90,000 perimetric examinations. A simulation model that based on the analysis was used to verify the results. The deviation from the age-corrected normal threshold of six points representative of six regions defined by Garway-Heath et al. (2000) and Gonzalez de la Rosa et al. (2002a) showed a comparable and high correlation coefficient with the mean deviation (MD) when original values were used as well as when the values obtained from the simulation model were used. By using all four phases to obtain the final average results, the standard error of MD was lower than the initial phase with 0.5 dB or less for all the cases and even lower for early glaucoma. Estimation of local deviation produced a reasonably low standard error as well. SPARK was shown theoretically by using information from a set of limited points; it can theoretically produce rapid and stable results (Gonzalez de la Rosa and Gonzalez-Hernandez, 2013). Averaging the thresholds obtained in the four phases of SPARK strategy can lead to a reduction in the fluctuation of about 40% in the early stages of glaucoma (Gonzalez de la Rosa and Gonzalez-Hernandez, 2011).

The idea of using a limited number of points which are selected according to their relation with glaucomatous damage was started by Krakau (1989). Nevertheless, there were some doubts about using the limited number of points may miss some focal defects and the study earlier was based on simulation models and retrospective analysis. A recent prospective observational case-control study which used the SPARK strategy in Oculus Easyfield perimeter to compare the diagnostic capacity of its first and final phases, and also three

procedures of glaucoma morphologic analysis using Heidelberg retinal tomography (HRT), Zeiss scanning laser polarimetry (GDx), and Cirrus optical coherence tomography (OCT) (Gonzalez de la Rosa et al., 2013). It used a sample of normal eyes and eyes with suspected and confirmed glaucoma which has at least two times perimetric experience. Mean deviation (MD) of especially second and final phase obtained the maximum diagnostic capacity compared to the rest of the morphology indices. In order to achieve 95% specificity, the optimum cut off value of final phase MD was -2.3 dB and PSD was 1.8 dB. The diagnostic capability of the first phase of SPARK was poor when the PSD was used. However, the first and final phase showed a very good kappa agreement. Both also showed about the same kappa agreement to the morphologic indices and no significant difference in their correlation coefficients to all the morphologic indices used in the study. SPARK could produce high sensitivity and specificity even with the first phase which takes less than 40 s but further study needed to determine which phase produces the most stable result in order to detect the progression of visual field defect.

The SPARK strategy was also used to evaluate the performance of the new method introduced for glaucoma detection i.e. the glaucoma discriminant function (GDF) of the Laguna ON_hE (optic nerve haemoglobin) program. The program is designed to calculate the haemoglobin (Hb) amount at the ONH which showed high sensitivity and strongly correlated to SPARK strategy (Gonzalez de la Rosa et al., 2013a). Their correlation was better than with rim sector areas or the corresponding nerve fibre thickness using OCT (Pena-Betancor et al., 2015). Both studies defined the abnormal VF using SPARK as reproducible glaucomatous VF loss without any other abnormalities that caused the defect.

The results of the SPARK strategy can also be used to estimate the morphologic defects in particularly its estimated threshold values which were used to deduce the thickness of RNFL. The values were used to derive a normalized value of the neuroretinal rim area, correcting the influence of the optic disc area size and expressing the result as a percentage of their normal

average relation. (Gonzalez de la Rosa et al., 2015) There were good correlations between the deduced and measured RNFL and also between the deduced and measured rim area. High inter-individual variability and also significant test-retest fluctuation were found in morphology. The morphologic estimates obtained from the SPARK results showed only slightly lower accuracy than the test-retest error of direct morphologic measurement procedures (Gonzalez de la Rosa et al., 2015).

A proper comparison of the global indices of SPARK and SS is yet to be found in peer-reviewed journals but the normal values of differential luminance sensitivity (DLS) of the Oculus Twinfield perimeter were reported to be 1.5 dB higher compared with HFA using full threshold strategy (Lorch et al., 2001). The reference luminance which is the maximal stimulus intensity that each instrument is capable of producing from of Oculus Twinfield Perimeter is different from HFA. Twinfield's value was reported as 2511 asb and HFA's was 10000 asb whereas their respective values of minimal stimulus luminance are 0.31 asb and 1.00 asb (Schiefer, et al., 2005). The current Oculus Twinfield can only produce maximum stimulus intensity physically up to 1000 asb but it runs SPARK with decibel scale according to simulated intensity up to 10,000 asb (3180 cd/m^2). Both instruments are using similar background luminance i.e. 10 cd/m^2. The comparison between the parameters used in HFA and Oculus Twinfield when performing VF test using strategy test, SITA and SPARK respectively is shown in Table 1.1.

Table 1.1: Comparison between HFA and Oculus Twinfield when using strategy test, SITA and SPARK respectively

Parameters	HFA	Oculus Twinfield
Testing distance	30 cm	30 cm
Bowl shape	aspherical	spherical
Background luminance	31.5 asb	31.4 asb
Background light source	Fluorescent light	Halogen light
Maximum stimulus luminance	10,000 asb	10,000 asb (Simulated)
Stimulus size/colour	Goldmann size III/white	Goldmann size III/white
Stimulus shape	round	round
Stimulus duration	200 ms	200 ms
Stimulus interval	1.5 – 4.0 s	1.6 s
Fixation control	Heijl Krakau (Blind spot)	Central threshold

1.5 Global Indices

There are a number of VF indices developed for Octopus (Flammer, 1986) and HFA (Heijl et al., 1987a) to provide single summary values or overview of the static perimetry result which could provide easier interpretation or description of the changes over different periods of VF results. As higher consistency was found with fast threshold strategy, the global indices that are more commonly used currently are mean deviation (MD) and pattern standard deviation (PSD).

1.5.1 Mean Sensitivity (MS)

Mean sensitivity (MS) is the least important index clinically which represents the average of the threshold values from all the testing points in the VF (Eq 1.1). It is a least manipulated index as age-matched normative data is not required.

$$MS = \frac{1}{m}\sum_{i=1}^{m} x_i \qquad \text{(Eq 1.1)}$$

Symbols	Meaning
x_i	sensitivity for test location *i*
m	number of stimulus location excludes the blind spot

(Flammer, 1986)

Mean sensitivity (MS) is provided in the result of SPARK in addition to the two indices but it is not displayed in SITA result. General reduction of sensitivity or diffuse loss of the VF can cause a reduction in MS. Shallow focal loss/scotoma may not cause a significant change in MS (Flammer, 1986). The average range of MS is age-dependent. Spry and Johnson (2001) reported that MS deteriorates with 0.43 dB/decade before 53.4 years old and 1.02 dB/decade after the age of 53.4 years old. Different types of perimetry exhibit different rates of age-related sensitivity loss (Gardiner et al., 2006). Mean sensitivity is not used clinically to determine whether the VF result is normal.

1.5.2 Mean Deviation (MD)

Generally, mean deviation summaries the overall changes of the threshold values from all test points which are relative to the age-adjusted normal population values (Flammer 1986). It is also a weighted average value in certain perimetry (Wall and Johnson, 2005; Heijl et al. 1992) such as HFA. The weight of each point is according to the magnitude of normal range at that point and it is weighted for eccentricity. Points that are nearer to central fixation are more important and less variation are weighted more heavily (Funkhouser and Fankhauser, 1991). It is one of the VF indices widely used in all threshold perimetries to provide summary statistics of the VF test result. The formula for MD is shown in Eq 1.2.

$$MD = [\frac{1}{m}\sum_{i=1}^{m}\frac{(x_i - z_i)}{S_{1i}^2}]:[\frac{1}{m}\sum_{i=1}^{m}\frac{1}{S_{1i}^2}] \qquad \text{(Eq 1.2)}$$

Symbols	Meaning
S_{1i}^2	variance of the normal field measurement at location *i*
z_i	normal reference threshold at location *i*
x_i	measured threshold of test location *i*
m	number of test locations (excluding the blind spot)

(Anderson and Patella, 1999)

Mean deviation (MD) is one of the global indices used in HFA. Negative values of MD indicate a loss in sensitivity. Its magnitude represents the degree of the generalised VF loss and high diffuse loss could be due to any VF defects. However, the presence of a large focal loss will cause a large number of locations to have depressed sensitivity, hence the negative value of MD will also be increased (Cubbidge, 2005). The reduction of MD (become more negative value) can be due to any pre-retinal opacities such as corneal lesion, cataract and other media opacities which indicate further diffuse field loss (Cubbidge, 2005). Poor reliability indices such as high false positive rate in SITA VF test could affect both MD and PSD (Newkirk et al., 2006).

Each threshold perimeter that is commercially available could have differences in the MD definition (Heijl et al., 1992; Flammer et al., 1985) and its collection of age-corrected database which explains the MD values produced by each perimetry are not directly comparable. The different instrument set up including background illumination and maximum stimulus intensity will also produce MD values that are not comparable between perimeters (Papp et al., 2001). Some perimetries even produce an inverted sign of the mean defect (MDf) such as in Octopus perimetry (Haag-Streit, Koeniz, Switzerland) and Henson perimetry (Elektron Technology) which is a non-weighted mean value. A positive MDf will be obtained if the age-corrected normal threshold value is higher than the measured threshold value at the test points which indicates sensitivity loss. Its formula is shown in Eq 1.3.

$$MD_{\text{Octopus}} = \frac{1}{n}\sum_{i=1}^{n}(z_i - \bar{x}_i) \qquad \text{(Eq 1.3)}$$

Symbol	Meaning
z_i	age-corrected normal value of test location *i*
$\bar{x}_i$	value of test location *i* (estimated as $\bar{x}$) if repeated measurements are available
n	number of test locations (excluding the blind spot)

(Flammer et al., 1985)

Nevertheless, the differences between MD and MDf were reported to be negligible which may be considered to be interchangeable using certain types of VFs or perimeters (Funkhouser and Fankhauser, 1991). The weighting function was shown to have little effect on the MD but it caused slight increases in PSD and CPSD (Flanagan et al., 1993b).

Mean deviation is not commonly used as the first evidence to confirm the detection or diagnosis of VF defect. Higher variability of MD was found in eyes with VF defect (Chauhan et al.,2008; Jampel et al.,2006; Heijl et al.,1989a). It was not included in the six highly specific criteria to detect glaucoma in a study conducted by Johnson and colleagues (Johnson et al., 2002). A widely used criteria recommended by Hodapp-Parrish-Anderson (HPA) (Hodapp et al., 1993) for the diagnosis of glaucoma using SITA strategy does not include MD but it was used in Octopus perimetry with MDf more than 2 dB is one of the diagnosis criteria (King et al., 2002; Wadood et al., 2002; Morales et al., 2000). A mean defect (MDf) of more than 2.6 dB is also used as one of the criteria to diagnose a glaucomatous VF defect in G1 program using Octopus perimeter (Horn et al., 2005). The value of the MD is used to determine the level of VF defect to normal, borderline or one of three grades of defect which is found in Octopus perimeter (Shelat and Rao, 2009). It was also used by Hodapp et al. (1993) as one of the criteria to classify the severity level of glaucoma which applied on the 24-2 test grid. One of the criteria to classify an eye as early glaucoma is its MD must be less than -6 dB but more than -12 dB. Moderate and severe glaucoma will be graded only if MD is lower than -12 dB.

SPARK strategy in Oculus perimeter (Oculus Optikgeräte GmbH, Wetzlar, Germany) uses a non-weighted average value of the differences between the estimated threshold value and its age-corrected normal value at all 66 test points for its MD value but it has inverted sign compared to mean defect (MDf) used in Octopus perimeters. It was defined interchangeably as mean defect or mean deviation by Gonzalez de la Rosa and colleagues (2013) but mean deviation is preferred and more frequently used as the mean value does not necessarily represent a defect of VF (Pena-Betancor et al., 2015; Gonzalez de la Rosa et al., 2015; Gonzalez de la Rosa et al., 2013a). A negative MD value in SPARK means a loss of sensitivity that is similar to the MD in SITA.

SPARK strategy uses a combination of limited test points that best correlates to MD to estimate the threshold values (Gonzalez de la Rosa and Gonzalez-Hernandez, 2013). By using MD of this strategy with a cut-off point of -2.3 dB, the sensitivity of 86.5% and specificity of 95.1% were achieved in the diagnosis of glaucoma patients which were comparable to morphologic analysis (Gonzalez de la Rosa et al., 2013).

1.5.3 Pattern Standard Deviation (PSD)

Pattern standard deviation (PSD) is another global index used in standard automated perimetry such as HFA. It is defined as a weighted standard deviation of the difference of the measured sensitivity value at each stimulus location compared to the age-corrected normal values. The weighting function used is the same as in the calculation of MD. In the measure of PSD, the difference of each measured sensitivity value from the normal value is weighted based on the variance of the normal values at the test location. Besides that, the calculation of PSD from SITA algorithm is taking into account the correction for short-term fluctuation (King et al., 2002). Its formula used in HFA is shown in Eq 1.4:

$$PSD = \sqrt{\left[\frac{1}{m}\sum_{i=1}^{m} S_{1i}^2\right]\left[\frac{1}{m-1}\sum_{i=1}^{m}\frac{(x_i - z_i - MD_{HFA})^2}{S_{1i}^2}\right]} \qquad \text{(Eq 1.4)}$$

Symbol	Meaning
MD_{HFA}	mean deviation as used in HFA
S_{1i}^2	variance of the normal field measurement at location *i*
z_i	normal reference threshold at location *i*
x_i	measured threshold of test location *i*
m	number of tested locations (excluding the blind spot)

(Anderson and Patella, 1999)

Pattern standard deviation represents the degree of irregularity of the VF sensitivity values from the normal hill of vision and it indicates the amount of localised VF loss (Wall and Johnson, 2005). Scotomas produce significant deviation from the normal value of the VF slope and increase the degree of the irregularity, hence increased PSD will be detected. Pattern standard deviation will only show positive value with a small value indicates a smooth uniform hill of vision, while an irregular hill of vision will produce a large PSD (Yaqub, 2012).

Pattern standard deviation is used to diagnose a patient with localised VF defect as it can quantify the amount of the focal loss and any progression in the early stage of glaucoma but it is not helpful in monitoring advanced glaucomatous defects (Chaglasian, 2013). Pattern standard deviation increases in the early stage of the field loss but declines at the advanced stages as almost the entire VF is affected with all stimulus points are equally defective without any significant localised defects (Broman et al., 2008). Pattern standard deviation increases as well as MD when there is a presence of profound localised defect (Gonzalez de la Rosa et al., 2013). A large value of MD with a normal value of PSD indicates diffuse VF loss with cataract is the most common cause. An increased MD but relatively constant PSD indicates a worsening cataract whereas if a stable MD with increased PSD, most likely the glaucoma is progressing. If both MD and PSD are abnormal, this may be due to either a generalized but non-homogenous defect or a large localized defect (Brusini and Johnson, 2007). Pattern standard deviation is very minimally affected by cataract.

Due to the PSD always shows the sign of increment during the early stage of glaucoma, it is widely used as one of the early evidence for glaucomatous VF loss. It is used as one of the guidelines recommended by HPA (Hodapp et al., 1993) based on HFA printout (30-2) to diagnose VF defect when PSD is significantly outside the normal range with p < 5%. In Ocular Hypertension Treatment Study (OHTS), abnormal corrected PSD (p < 0.5%) was used alongside with GHT as the criteria to detect the presence of the glaucomatous VF loss (Keltner et al., 2000; Gordon and Kass, 1999; Johnson et al., 2002). Corrected PSD is one of the global indices found using FT in SAP. It is adjusted according to short-term fluctuation. It is not found in the fast threshold strategy. Pattern standard deviation with p<1% is one of the criteria that produced high specificity (98 – 100% for SAP) to identify glaucomatous eye (Johnson et al., 2002a). It also has the largest area under the receiver operating characteristic curve (AUC) among the global indices for detecting the earliest evidence of glaucomatous damage (Goldbaum et al., 2002). Pattern standard deviation also performed the best using standard achromatic automated perimeter in determining the progressive glaucomatous optic neuropathy which is comparable to the criterion using the number of test locations demonstrating p<1% on total deviation plot (Sample et al., 2006). Generally, PSD is considered not useful after the diagnosis of the glaucoma is confirmed. The larger mean of PSD was found to be significantly correlated to the progression of VF loss in better eyes (Chen, 2002). Higher PSD values from Humphrey automated perimeter at baseline was one of the predictive factors for the development of open-angle glaucoma in patients with ocular hypertension reported in The Ocular Hypertension Treatment Study (OHTS) (Gordon et al., 2002) and European Glaucoma Prevention Study (Miglior et al., 2007).

Pattern standard deviation is not used in certain standard automated perimeter such as Octopus perimeter in which loss variance (LV) is employed as a global index instead of PSD. Its formula is shown in Eq 1.5.

$$LV = \frac{1}{m-1}\sum_{i=1}^{m}\left(z_i - \bar{x}_i - MD_{Octopus}\right)^2 \qquad \text{(Eq 1.5)}$$

Symbol	Meaning
$MD_{Octopus}$	mean defect as used in Octopus
z_i	age-corrected normal value of test location *i*
$\bar{x}_i$	value of test location *i* (estimated as $\bar{x}$) if repeated measurement are available
m	number of tested locations

(Flammer et al., 1985)

The weighting factor is not applied in this formula and it is called standard deviation defect in Henson perimeters. Both indices are analogue to PSD in HFA (Cubbidge, 2012).

Loss variance is used as one of the criteria for the diagnosis of abnormal VF. As recommended by the manufacturer of the Octopus perimeter (Octopus, 1991), LV > 6 dB was used as one of the criteria in a few studies using Octopus perimeter (King et al., 2002; Morales et al., 2000). Loss variance is used similarly to PSD as a measure of the degree of focal VF loss typically for glaucoma. It presented a very high correlation with PSD which indicates both could be used to identify a similar degree of focal VF loss (King et al., 2002).

Pattern standard deviation is employed in SPARK strategy using a simpler formula without weighting factor employed by SITA (Eq 1.6). Nevertheless, due to both strategies are using slightly different background illumination and individual age-matched normal values, both indices are not interchangeable.

$$PSD = \sqrt{\left[\frac{1}{m-1}\sum_{i=1}^{m}(x_i - z_i - MD_{Oculus})^2\right]} \qquad \text{(Eq 1.6)}$$

Symbol	Meaning
MD_{Oculus}	mean deviation as used in Oculus
z_i	normal reference threshold at location *i*
x_i	measured threshold of test location *i*
m	number of tested locations (excluding the blind spot)

The optimum cut-off point of PSD using SPARK is 1.8 dB with specificity of 95.1% and sensitivity of 82.7% were obtained in glaucoma diagnosis and a good agreement with morphologic indices was reported by Gonzalez de la Rosa et al. (2013). Pattern standard deviation that obtained through the first phase in SPARK is relatively poorer in diagnostic capacity but its diagnostic capacity is much improved in the final phase of SPARK (Gonzalez de la Rosa et al., 2013).

1.6 Glaucoma in Threshold Perimetry

Glaucoma is regarded as a progressive optic neuropathy associated with continuous degeneration of retinal ganglion cells which exhibits typical structural changes of the optic nerve and characteristic VF loss (Otarola et al., 2016; Weinreb and Khaw, 2004; Fechtner and Weinreb, 1994). As the visual function deterioration is irreversible and always asymptomatic in the early stage, blindness is a possible ultimate consequence (Hattenhauer et al., 1998; Oliver et al., 2002). It is one of the leading causes of blindness (Flaxman et al., 2017; Bourne et al., 2013; Tham et al., 2014; Congdon et al., 2004; Pascolini et al., 2004; Resnikoff et al., 2004) which affects approximately more than 60 million people worldwide in 2010 and potentially reaching 79.6 million by 2020 with 11.2 million of them are blind bilaterally (Quigley and Broman, 2006). It is also estimated that over 3 million populations in the United States will be affected by open-angle glaucoma by 2020 (Friedman et al., 2004). More than half of the glaucoma patients do not even realise that they have the disease (Tielsch et al., 1991; Quigley, 1996; de Voogd et al., 2005) and up to 24% of them were blind at least unilaterally before end of their lives (Mokhles et al., 2016; Peters et al., 2013; Forsman et al., 2007).

Glaucoma was first introduced as a word used to describe a dimming of vision in the elderly by Hippocrates in Greece in approximately 400BC (Leffler et al., 2015). But it was only known to be associated with elevated intraocular pressure since 1622 by English ophthalmologist Richard Bannister (Realini, 2011; Grewe, 1986) and the physically glaucomatous changes in the fundus were only able to be observed after the invention of ophthalmoscopy in the 1850s.

It was only discovered a decade later that high intraocular pressure (IOP) plays a role in causing blindness without inflammation when “Simple Glaucoma” was introduced. It was first defined as a disease of the optic nerve which links to a number of risk factors in 1973.

Some examples of characteristic VF defects by glaucoma were published by Albrecht von Graefe in the 1850s but at that time, it was regarded as amblyopia and the relationship between optic disc cupping, increased IOP and VF loss was not fully established yet (Johnson et al., 2011). With the evolving of the quantitative perimetry from Bjerrum screen to Goldmann perimetry and further to standard automated perimetry, VF test has developed to be one of the important procedures to evaluate the visual function of glaucoma patients (Jampel et al., 2011). Automated static perimetry was capable of detecting the glaucomatous field defects one year earlier than manual perimetry with the combination of kinetic and static stimulus in three-quarter of the ocular hypertension patients who had developed glaucomatous field loss (Katz et al., 1995). It has become the benchmark for testing visual function in glaucoma. The use of information related to glaucomatous VF defects in various fast threshold strategies such as SITA and ZATA helps to shorten the testing time with minimal compromise of the sensitivity and specificity of the test (Bengtsson and Heijl, 1998; Harvey, 2011). The introduction of such fast threshold strategy has made the threshold perimetry clinical routinely used to determine the extent of glaucomatous damage to visual function and the progression of visual loss. It was once regarded as the gold standard for glaucoma diagnosis (Dada and Mandal, 2008; Wood et al., 2000).

Visual field loss is one of the hallmarks of glaucoma (Artes and Chauhan, 2005; Drance, 1972; Gliklich et al., 1989; Hart and Becker, 1982). It correlates to its glaucomatous optic nerve appearance. Paracentral scotomas, Seidel scotomas, arcuate scotomas, isolated nasal steps and altitudinal scotomas are among the classic glaucomatous field defects (Sihota et al., 2007; Steele and Spry, 2009). The patterns of glaucomatous field defect correspond to the routes of the retinal nerve fibre which meet along the median raphe and central vision is usually not

affected (Sihota et al., 2007; Lau et al., 2003; Hoffman et al., 2006; Sample et al., 2004; Goldbaum et al., 2005). Arcuate scotomas and nasal steps are among the common field defects found in early glaucoma (Risse et al., 1999; Hart and Becker, 1982; Lee et al., 2003). Arcuate scotomas start from the blind spot and extend over or under the central fixation, or both, to the horizontal median raphe which is correspondent to the arcuate retinal nerve fibres. Meanwhile, nasal steps always start as step-like localized scotomas in the nasal periphery where the nerve fibres meet along the median raphe (Werner and Beraskow, 1979). Different types of glaucoma have different spatial patterns of glaucomatous field defects and structural properties of the optic disc (Rhee et al., 2001; Gazzard et al., 2002; Lau et al., 2003; Boland et al., 2008; Nouri-Mahdavi et al., 2011). Two different mechanisms hypothesized in glaucomatous optic neuropathy such as pressure-dependent and pressure-independent mechanisms may play a role in the differences of the patterns (Rhee et al., 2001; Gazzard et al., 2002; Nouri-Mahdavi et al., 2011). Normal-tension glaucoma (NTG) which is less pressure-dependent, exhibited VF defects that were localized, closer to fixation, with a steeper slope and greater depth (Caprioli and Spaeth, 1984; Araie et al., 1993). Whereas primary angle-closure glaucoma (PACG), regarded as the more pressure-dependent glaucoma was found to have more localised damage in superior hemifield than inferior hemifield (Gazzard et al., 2003; Atalay et al., 2016) and it was more obvious in more advanced glaucoma (Atalay et al., 2016). Paracentral scotomas were less likely to be found in chronic angle-closure glaucoma (CACG) compared to primary open-angle glaucoma (POAG) (Nouri-Mahdavi et al., 2011a). A more diffuse sensitivity reduction in POAG eyes with higher IOP was found compared to POAG eyes with lower IOP which had more localised damage (Caprioli et al., 1987; Chauhan et al., 1989). Nevertheless, there are other studies showing no difference between NTG and POAG (Araie et al., 1993; Motolko et al., 1982; Lewis et al., 1983; King et al., 1986; Caprioli et al., 1987). The inconsistencies in defining and classifying POAG and NTG patients across the studies could attribute to the variable findings (Atalay et al., 2016).

Mild generalized depression of field sensitivity (Anctil and Anderson, 1984; Stamper, 1984; Drance, 1991; Lachenmayr et al., 1991; Corallo et al., 1995; Lachenmayr et al., 1992; Henson et al., 1999; Polo et al., 2002; Kitazawa and Yamamoto, 1997; Gonzalez de la Rosa et al., 2010) was also reported in the early stage of glaucoma but it is arguable that many factors such as media opacity, miosis, or retinal dysfunction could also be the cause of the diffuse loss (Anctil and Anderson, 1984; Stamper, 1984; Drance, 1991, Lachenmayr et al., 1992; Chauhan et al., 1997; Henson et al., 1999; Heijl, 1989; Asman and Heijl, 1994; Langerhorst et al., 1989). As such, it is not preferably used for the diagnosis of glaucoma. Central VF defects within 10 degrees of fixation were also reported in early glaucoma (Hood et al., 2011; Traynis et al., 2014; Shafi et al., 2011) but 10-2 field loss was rare in subjects with normal 24-2 and spectral-domain optical coherence tomography results (Sullivan-Mee et al., 2016).

Peripheral field defects could be the only detectable signs in some patients suffering from early glaucomatous field defects (LeBlanc and Becker, 1971; Armaly, 1971; Werner and Beraskow, 1979). Individuals with a normal central field (24° - 30°) could have a manifest abnormal peripheral field (Odden et al., 2016). More peripheral than central field loss can be found in the early stage of glaucoma. In fact, it is also possible to have more peripheral field sparing than the central field in advanced glaucoma (Odden et al., 2016). It was reported long ago that the spared vision in advanced glaucoma was regarded as “temporal island” (Posner and Schlossman, 1948) and Brais and Drance (1972) had suggested that peripheral testing should be performed in advanced glaucoma patients with an extensive central loss to monitor the glaucomatous progression. Nevertheless, it was reported that only 4% - 11% of glaucoma or suspected glaucoma eyes with normal central VF demonstrate peripheral field defects (outside of central 30° field) (Miller et al., 1989; Ballon et al., 1992; Caprioli and Spaeth, 1985; Stewart et al., 1988; LeBlanc and Becker, 1971). There is a higher potential to have the nose and/or eyelid artefacts in peripheral field testing especially in patients with a large nose and/or deep-set orbits (Gramer et al., 1982). Greater variability was exhibited in the peripheral field especially in the superior and nasal quadrants (Young et al., 1990; Stewart and Shields, 1991)

therefore it is more difficult to determine the normal value for the peripheral test points (Odden et al., 2016) even though some expected the hill of vision on the peripheral has been adjusted according to the eccentricity, pupil and uncorrected refractive error (Odden et al., 2016; Young et al., 1990). More false positive may be obtained during the peripheral VF test which could affect the diagnosis (Caprioli and Spaeth, 1985). Peripheral field defects also could be normalized after repeated testing (Stewart and Shields, 1991). Peripheral field assessments also do not provide additional information for the determination of glaucoma progression if central field damages are present (Schulzer et al., 1987; Enoch, 1978). In fact, it is a common practice that only central 24° to 30° of VF is measured with automated static perimetry in the diagnosis and management of glaucoma (Stewart and Sheilds, 1991; Freeman et al., 2007). The increased time required in the threshold test must be justifiable for the added information in the quantitative testing despite peripheral nasal VF was shown to provide valuable additional information whereas temporal peripheral field testing was not (Seamone et al., 1988). The stimuli located in superior and inferior arcuate were shown to be the most informative whereas stimuli in the peripheral superior field (> 20°) and around blind spot were least informative. The use of these more informative test points helped to increase the diagnostic power and at the same time shorten the testing time (Henson et al., 1984; Henson and Chauhan, 1985). It is not necessary to achieve high sensitivity of the test with a large number of test points (Henson et al., 1988). It is better to run the test in a shorter time as test time below 3 minutes could reduce the variability of the patient's response (Henson and Emuh, 2010). It was estimated that around 66% of retinal ganglion cells are responsible for the central VF and 83% of the striate cortex is occupied by the central field of 30° to 40° (Yaqub, 2012). There is no validated and standardised method in documenting the glaucomatous peripheral field loss (Odden et al., 2016). Central VF testing was shown numerous times in detecting early signs of glaucoma with reasonably shorter testing time which indeed minimizes the fatigue effects (Aulhorn and Karmeyer, 1977; Gramer et al., 1982; Heijl and Lundqvist, 1984; Heijl, 1989a; Heijl et al., 2012). With the introduction of automated static perimetry, Central 30° or 24° threshold perimetry has been a benchmark of visual function assessment in glaucoma (Susanna and Vessani, 2009;

Jampel et al., 2011). It was used widely in all the major clinical trials in glaucoma such as the Collaborative Normal-tension Glaucoma Study (CNTGS) (Drance et al., 2001; CNTGS group, 1998; 1998a), the Advanced Glaucoma Intervention Study (AGIS) (The AGIS investigators, 1994; Ederer et al., 1994), the Collaborative Initial Glaucoma Treatment Study (CIGTS) (Musch et al., 1999; Gillespie et al., 2003), the Ocular Hypertension Treatment Study (OHTS) (Gordon et al., 1999; Gordon et al., 2002), the Early Manifest Glaucoma Treatment Study (EMGT) (Leske et al., 1999) and the European Glaucoma Prevention Study (EGPS) (Miglior et al., 2002; Miglior et al., 2005). Full threshold strategy test was used in all these clinical trials but different criteria were used to diagnose and detect the progression of glaucoma such as scores based on deviation from total deviation plot, changes in pattern deviation, the value of PSD and Glaucoma Hemifield test (GHT). Computerized threshold perimetry has been equipped with the specifically designed analytical program that can detect glaucomatous VF abnormalities such as general reduction of sensitivity, localized defects and asymmetrical sensitivities between vertical hemifields (Sommer et al., 1985; Sommer et al., 1987; Flammer, 1986). Moreover, the VF global indices (MD and PSD) can even provide quantitative information for the severity of the defects. The use of pattern deviation probability map further differentiates the localized defects which might be masked by media opacities (Bengtsson et al., 1997a).

Glaucoma Hemifield test (GHT) is a program incorporated in HFA Statpac to detect asymmetry between superior and inferior hemifields. It is one of the criteria to determine the endpoint of the clinical trial in OHTS (Keltner et al., 2000) and as an eligible criterion used in EMGT (Leske et al., 1999). It compares the threshold values in mirrored sectors of superior and inferior hemispheres (Duggan et al., 1985; Asman and Heijl, 1992a; 1992; Sommer et al., 1987). Five corresponding sectors are chosen based on the course of retinal nerve fibre bundles and are analyzed by GHT. A score is assigned to each test point according to the value shown in the pattern deviation probability map and the total score of each sector is calculated. A large normal database is used to determine the significance of the differences between scores of

both mirrored sectors. “Outside normal limits” will be indicated by GHT if the difference of any mirrored sectors is more than 99.5% of normal limits. “Borderline” will be classified if any sector differences are within 97% to 99.5% of normal limits. If all the sectors do not show differences more than these two limits, “within normal limits” is categorized for the field (Asman and Heijl, 1992; Ghazali et al., 2015). Glaucoma Hemifield test has produced high sensitivity and specificity in detecting glaucomatous VF loss (Asman and Heijl, 1992; Katz et al., 1991; Sample et al., 2006; Johnson et al., 2002a; Susanna et al., 1994). It has better ability in discriminating glaucoma from the normal field compared to using global indices (MD and PSD) with its specificity increased in the repeated test (Katz et al., 1995a; Katz et al., 1996). Glaucoma Hemifield test can enhance the detection of early VF defects but it is not only used specifically for glaucoma (Nouri-Mahdavi et al., 2011).

Glaucoma Hemifield test with “outside normal limits” is one of the criteria used in HPA criteria that are used widely to determine the cut-off point of glaucoma diagnosis or the earliest evidence of glaucomatous field loss. It is based on the central 30-2 to define a VF defect (Hodapp et al., 1993). The criteria also include PSD with value only found in less than 5% of the normal population and a hemifield cluster of minimum 3 non-edged points on pattern deviation plot at $p < 5\%$ with at least one point of them showed $p < 1\ \%$. Hodapp-Parrish-Anderson (HPA) criteria are unlikely to have a false positive result for a patient with VF defect (Chakravarti, 2017).

Johnson et al. (2002a) showed that there are a few criteria found to produce high specificity in the diagnosis of glaucoma besides GHT. They include i) PSD with $p < 1\%$; ii) One hemifield cluster demonstrating $p < 1\%$; iii) Two hemifield clusters with $p < 5\%$; iv) presence of four abnormal test locations ($p < 5\%$) on pattern deviation probability plot; v) five abnormal test locations ($p < 5\%$) on pattern deviation probability plot. Sample et al. (2006) had also shown that PSD and a criterion based on the number of test locations demonstrating $p < 0.1\%$ on

total deviation plot were the best diagnostic criteria in progressive glaucomatous optic neuropathy.

The threshold variability of glaucomatous field is larger (Heijl 1977; Holmin and Krakau, 1979; Flammer et al., 1984) and variability increases when the retinal sensitivity reduces (Wilensky and Joondeph, 1984; Lewis et al., 1986; Heijl et al., 1987; Gardiner et al., 2014). The variability found in the glaucoma patients was also higher compared to normal subjects (Chauhan et al., 1993) and more advanced field defect produced higher test-retest variability (Artes et al., 2005; Blumenthal et al., 2000; Gardiner et al., 2012; Henson et al., 2000; Jampel et al., 2006). A higher number of VF examinations is required to achieve statistical significance of glaucomatous field defect. It was typically shown in EMGT (Heijl et al., 2003) and OHTS (Keltner et al., 2006) where three consecutive field defects were used as an endpoint. Chauhan et al. (2008) have recommended that a glaucoma patient should perform six VF examinations in the first 2 years to obtain a good baseline data. The advancement in the glaucomatous structural test further improves the glaucoma diagnostic ability when combined with the VF tests (Shah et al., 2006; Mardin et al., 2006; Hong et al., 2007). Even though structural test such as OCT-measured RNFL defects has an advantage of early glaucoma detection with its low intersubject variability (Harwerth et al., 2007) but standard automated perimetry is still heavily reliant on for the glaucoma diagnosis and progression (Nouri-Mahdavi, 2014; Camp and Weinreb, 2017) especially for the moderate and advanced stages of glaucoma. Introduction of non-conventional perimetry such as short-wavelength automated perimetry (SWAP) (Johnson et al., 1993b; 1993a; Demirel and Johnson, 2001; Soliman et al., 2002) and frequency doubling technology (FDT) (Johnson and Samuels, 1997; Medeiros et al., 2006; Leeprechanon et al., 2007) initially showed superiority in the detection of early glaucoma compared to achromatic standard automated perimetry. But the performance of SWAP is limited by cataractous eyes, high test-retest variability, long test duration and additional learning effect (Sample, 2001; Blumenthal et al., 2003; Wild, 2001) whereas FDT has a limitation in the detection of glaucoma progression (Redmond et al., 2013a; Sample et al.,

2000). Both perimetries have yet to be used frequently for glaucoma monitoring in clinical practice (Fogagnolo et al., 2008).

1.6.1 Advanced Glaucoma Intervention Study (AGIS) Scoring System

Advanced Glaucoma Intervention Study (AGIS) is a multicentre randomized clinical trial which was set up since 1988 (Ederer et al., 1994). Its objective is to evaluate the outcomes of the two surgical treatments sequences using argon laser trabeculoplasty (ALT) and trabeculectomy when medical treatment alone no longer adequately controls glaucoma. An objective quantitative scoring system was developed in this study to assess the severity of the glaucomatous VF defects by using a single index (The AGIS investigators, 1994). This AGIS scoring system uses HFA 24-2 threshold test. It consists of score ranges from 0 (no field loss) to maximum 20 which indicates total field loss. It is based on the number and depth of threshold depression from age-matched normal values in STATPAC-2 total deviation plot of the HFA. The 24-2 test field is divided into three sectors: nasal area, the upper and lower hemifield remaining areas. Each sector contains areas with specific minimum depressions from the age-matched normal threshold values in order to be considered as abnormal. The minimum depressed value of the abnormal test points varies from 5 to 9 dB depending on the test locations. Such values are larger in the periphery than the centre and the superior field has higher values than the inferior. A maximum score of 9 can be achieved in each hemifield and 2 in the nasal area. The present of the nasal defect or nasal step will add the score in this AGIS system, besides the number and depth of abnormal test points within the hemifield defect (The AGIS investigators, 1994). An increase score of 4 or more from the baseline reference field on three consecutive tests indicates a progression of field loss in AGIS. This has shown adequate reliability in detecting the progression of glaucomatous VF defects in most of the cases by the AGIS investigators (1998). Besides that, an AGIS score of 19 or 20 which is found in three consecutive 6-month follow up visits is also a criterion of VF worsening (Kim et al., 2004). A glaucoma patient must have VF defect with AGIS score of minimum 1 but lower than 17 to be eligible for AGIS (The AGIS investigators, 1994; Kim et al., 2004).

The AGIS scoring system is used in glaucomatous patients to monitor the VF follow-up tests over time but it is not designed for clinical use as it is time-consuming and difficult to use. It is useful in grading the severity of the glaucomatous field defects in a standardized method which is important in clinical research (Brusini and Johnson, 2007). It is also considered to be conservative as the progression of the glaucoma was underestimated by AGIS compared to scoring systems used in other studies such as Collaborative Initial Glaucoma Treatment Study (CIGTS) and Early Manifest Glaucoma Treatment (EMGT) (Heijl et al., 2008; Nouri-Mahdavi et al., 2007; Nouri-Mahdavi et al., 2005; Katz et al., 1997; Katz, 1999; Katz et al., 1999; Vesti et al., 2003). But the conservativeness of the AGIS scores only appeared after 4 to 5 years of follow up and it detected more progression than pointwise linear regression (PLR) and Glaucoma Change Probability Analysis (GCPA) in the first 4 years of follow up (Nouri-Mahdavi et al., 2007). Once the progression of the glaucoma is shown by AGIS score, it is highly possible that the VF deteriorates (Nouri-Mahdavi et al., 2007; Kim et al., 2004).

The AGIS scoring system has a low false positive progression rate (The AGIS investigators, 1994) and showed good specificity comparable to CIGTS scores (Heijl et al., 2008; Vesti et al., 2003) and even better than EMGT (Heijl et al., 2008). The Early Manifest Glaucoma Treatment scoring system tends to detect progression more often and possibly one year earlier than AGIS and CIGTS (Heijl et al., 2008). There is no gold standard to detect progression in VF loss as it is unable to determine which system can identify the ‘real’ progression. Collaborative Initial Glaucoma Treatment Study (Gillespie et al., 2003) uses different criteria to determine progression compared to AGIS which makes its scores slightly higher than AGIS scores regularly (Katz, 1999) and it is slightly more sensitive than AGIS with the detection of the confirmed progression earlier by 1.4 to 2.3 years (Vesti et al., 2003). Brusini and Johnson (2007) have pointed out that AGIS scoring system fails to include the mild diffuse loss found in early glaucoma stage even though it is accurate based on the localised defects. There are also many other studies that used or compared with AGIS scoring system which is widely used

as a global event analyses of glaucoma progression (Girkin et al., 2000; Patel et al., 2007; Nouri-Mahdavi et al., 2004; Ng et al., 2012; Chen, 2002; Brusini and Filacorda, 2006; Lin et al., 2003; Kang et al., 2015).

Budenz et al. (2002a) were using AGIS scores for the comparison between SITA strategies and Full Threshold in HFA. The scores were more than 1.0 point higher in results using Full Threshold than SITA strategies (SF and SS). These could be due to AGIS scoring system using total deviation plot which provides lower MD in SITA strategies than Full Threshold (Wild et al., 1999; Bengtsson and Heijl, 1999a; Sharma et al., 2000; Heijl et al., 2000; Shirato et al., 1999).

1.7 Relationship of Cataract with Threshold Perimetry

Cataract is described as the opacity of the crystalline lens which is commonly found among the elderly patients as its prevalence increases with age (Vashist et al., 2011; Song et al., 2018). It is highly possible that cataract can be found in a glaucoma patient as the treatment of glaucoma may increase the risk of cataract formation such as trabeculectomy (The AGIS investigators, 2001; Lichter et al., 2001), trabeculoplasty and medications (Heijl et al., 2002). In addition, the prevalence of glaucoma also increases with age (Bourne et al., 2016; Kapetanakis et al., 2016; Mitchell et al., 1996; Salmon et al., 1993; Dielemans et al., 1994; Leske et al., 1994; Tielsch et al., 1991).

Cataract causes dimmer retinal illumination, blurred retinal image and also scattering of light in the eye (Zuckerman et al., 1973). The increased intraocular stray light in the cataract eyes (de Wit et al., 2006) which is mainly forward scattered light (toward retina) (van den Berg, 1995; de Waard et al., 1992) also causes the reduction of contrast sensitivity and disability glare (Smith, 2002; van den Berg et al., 2009). Minimal of scattering light can be seen in the early stage of cataract and it may not have any significant effect on visual acuity but threshold

measurement could be affected (Heuer et al.,1988). The scattered light affects the threshold measurement more than the light absorption (Bettelheim and Chylack 1985).

It is commonly known that media opacities affect VF test (Guthauser et al., 1987; Wood et al., 1989; Dengler-Harles et al., 1990). Generally, a diffuse and uniformly depressed sensitivity is found in cataractous eye (Lam et al., 1991; Guthauser and Flammer 1988; Budenz et al., 1993). It is due to the equally reduction of the stimuli brightness and background illumination by the opacity of the crystalline lens which displays changes in global indices and total deviation plot (Budenz et al.,1993). Increased variability in VF was also found in cataractous eyes (Carrillo et al., 2005). A significant improvement of mean deviation after cataract surgery in glaucoma patients was reported (Rehman Siddiqui et al., 2007; Koucheki et al., 2004; Smith et al., 1997; Hayashi et al., 2001; Kook et al., 2004; Musch et al., 2006; Ang et al., 2010; Kim et al., 2001) but there were also reports shown negligible effect of cataract on VF (Stewart et al., 1995; Carrillo et al., 2005) which could be due to only mild to moderate cataracts were included in the study. Larger changes were found in more advanced glaucoma (Smith et al., 1997; Guthauser and Flammer, 1988). Contrarily, Chen and Budenz (1998) showed mean deviation usually improved after cataract extraction in mild or moderate glaucoma but the improvement was not necessarily observed in eyes with the advanced stage of glaucoma. Another global index, PSD or corrected pattern standard deviation (CPSD) was mostly reported of worsening after cataract extraction (Stewart et al., 1995; Smith et al., 1997; Hayashi et al., 2001; Koucheki et al., 2004; Siddiqui et al., 2005; Rao et al., 2013; Gillies and Brooks, 1998). PSD could underestimate the severity of advanced glaucoma (Koucheki et al., 2004). As hypothesized by Hayashi et al. (2001), shallow scotomas may be improved after cataract extraction but the deep field defects are relatively unchanged and could cause increased PSD. But studies also showed unchanged PSD after cataract removal (Lam et al., 1991; Vijaya et al., 2005; Ang et al., 2010; Rehman Siddiqui et al., 2007). The significant poorer pointwise Total Deviation (Carrillo et al., 2005) and pointwise Pattern Deviation (Koucheki et al., 2004) were also reported after cataract extraction. Whereas Visual Field Index (VFI) remained

unchanged after the cataract surgery (Rao et al., 2013). AGIS score was reported to improve after cataract removal in patients participating in AGIS (The AGIS investigators, 2000). Pattern deviation maps are more helpful in differentiating the field defects caused by cataract or glaucoma as the maps were less affected by the cataract surgery compared to total deviation maps (Bengtsson et al., 1997a). Loss variance (LV) and Corrected Loss Variance (CLV) in Octopus also help in detecting glaucomatous field defects in a cataractous eye (Novak-Laus et al., 2007) but mainly in mild or moderate and not severe cases of glaucoma (Pearson et al., 1990).

Increased straylight in an eye with cataract affects the results of standard automated perimetry (SAP) (Guthauser and Flammer, 1988; Lam et al., 1991; Budenz et al., 1993), SWAP (Moss and Wild, 1994; Moss et al., 1995; Kim et. al., 2001) and frequency-doubling technology perimetry (FDT) (Artes et. al., 2003; Tanna et. al., 2004; Casson and James, 2006). The results of SAP and SWAP were more influenced by cataract-stimulating filters compared to FDT and also grating-resolution perimetry (Anderson et al., 2009). Bergin et al. (2011) also showed that significant changes in threshold estimates using SAP, FDT and flicker-defined perimetry (FDF) when moderate to large increases of intraocular straylight presented in the eye.

The effect of cataract on the VF results also depends on the severity and the position of the lens opacity. De Waard et al. (1992) found more scattered light in posterior subcapsular (PSC) cataract compared to cortical cataract and nuclear sclerosis. More intraocular straylight and poorer contrast sensitivity were found in eyes with PSC and nuclear-cortical cataract than nuclear or cortical cataract (Bal et al., 2011). Chung et al. (2016) showed VF result was least affected by nuclear cataract whereas glaucomatous eyes with posterior subcapsular cataract improved significantly in MD but not PSD or VFI after cataract removal if their MDs were better than -12dB. PSC patients had been significantly easier to miss some testing points (Casson and James, 2006). The effect of PSC is larger (Yao and Flammer, 1993) due to its location is closer to the nodal point of the eye (Baraldi et al., 1987). Cortical cataract affected the changes

in MD, PSD and VFI for the patients with early glaucoma (MD > -6 dB) (Chung et al., 2016). Underestimated PSD could be detected in eyes with advanced glaucoma and nuclear cataract (Chung et al., 2016). In eyes with nuclear cataract, central retinal sensitivity was depressed more than peripheral sensitivity using either small or large targets but other types of cataract depressed more central sensitivity if using small targets and more peripheral sensitivity using large targets (Wood, 1989). Blue-on-yellow perimetry was reported to be more easily affected by PSC while white-on-white would be influenced more by anterior cortical opacity (Moss et al., 1995).

Co-existence of cataract and glaucoma increases the difficulty in assessing the progression of glaucoma (Heider et al., 1991; Smith et al., 1997; Chen and Budenz, 1998; Hayashi et al., 2001). Pattern deviation map is recognised as a tool which is widely used to differentiate the changes caused by glaucoma progression or media opacity (Bengtsson et al., 1997a; Katz, 2000; Rehman Siddiqui et al., 2007).

CHAPTER 2

RATIONALE AND DESCRIPTION OF THE RESEARCH

Visual field testing undeniably is an important tool in visual function assessment in particularly in glaucoma (Phu et al., 2017; Jampel et al., 2011; Fingeret, 2016; Nouri-Mahdavi, 2014) but it entirely relies on patients' subjective responses demanding their concentration throughout the test. Threshold perimetry requires multiple testing of each test point to determine the threshold level. Thus, the test duration becomes inevitably time-consuming which could exacerbate the fluctuation and unreliability of the results (Johnson et al., 1988; Heijl and Drance, 1983; Gonzalez de la Rosa and Pareja, 1997; Hudson et al., 1994). The unstable results could delay the decision-making of the clinicians while prolonged and repeated VF tests frustrate the patients. Thus, VF testing is generally not well-accepted by both clinicians and patients. Therefore, shorter testing time was seen as the way forward and various threshold algorithms were introduced in recent years aiming to reduce the test duration without loss of accuracy in the sensitivity estimation. To date, the Swedish Interactive Threshold Algorithm (SITA) is well-recognized universally in achieving this objective and regarded as "The Gold Standard" by many clinicians for VF testing.

2.1 Rationale

A newer fast threshold strategy, SPARK is incorporated in Oculus Perimeters (Oculus, Wetzlar, Germany). It was introduced to improve stability of the VF results through the information from a limited set of test points which are capable to produce reliable threshold estimates within a short examination time (Gonzalez de la Rosa and Gonzalez-Hernandez, 2013). This thesis describes an extensive comparison between SPARK and SITA through four studies that were undertaken at SEGi University, in Petaling Jaya, Malaysia from 2016 to 2018. In the thesis, SPARK Precision (SP) was the threshold strategy used for SPARK to compare with SITA Standard (SS) from SITA strategy. The first study comprised of the between-visit

within-strategy analysis in three different groups of subjects i.e. glaucoma, cataract and normal control subjects. The second study was to determine the validity of the SPARK in the estimation of normal sensitivity. The third and fourth studies evaluated the performance of SPARK in assessing the VF defects against SITA in subjects with glaucoma and cataract respectively. Glaucoma and cataract groups were chosen so as to have a group with focal (glaucoma) and another one with diffuse (cataract) VF loss.

2.1.1 Between-visit and Within-strategy Threshold Estimates Analysis for SPARK Precision and SITA Standard

As SPARK claimed to reduce fluctuation of the VF results, this study was to evaluate the reproducibility of the threshold estimates by SP as compared to SS in three different groups representing no VF loss, focal loss and diffuse loss. It was aimed to illustrate the necessity of the initial visit for the subjects to minimize the learning effect before comparison between-strategy was conducted in the following chapters. The effect of the perimetric experience on the DLS estimation was evaluated between and within the visits for each strategy which could affect the outcome of the strategy comparison. The consistency of the strategy was also further evaluated through analysis of pointwise variation between the visits without the data of glaucoma and cataract patients that comprised of various types of VF defect which could complicate the analysis. In view of the global indices could be affected by high reliability errors in threshold estimation (Junoy Montolio et al., 2012; Newkirk et al., 2006) and the delay of the decision-making followed by repeated test, the methods used in monitoring reliability for each perimeter is important to determine the quality of the results. Therefore, the evaluation of the reliability indices in each strategy was also carried out in this chapter.

2.1.2 Between-strategy Comparison between SPARK Precision and SITA in Normal Subjects

Comparison between two VF strategies was conducted in normal subjects for this study as it is the first logical step to evaluate a new perimetric test (Bengtsson et al., 1998). Normal

subjects usually were found to have lower inter- and intra-test variability (Chauhan et al., 1993; Flammer et al., 1984a), less visual fatigue (Suzumura, 1988) and shorter VF testing time which indeed often results in narrower confidence levels that could easily detect smaller differences between the strategies statistically. Furthermore, the threshold estimate is the least manipulated index compared to other global indices and the effect of learning was exhibited in the first study with evidence of more reliable threshold sensitivities in the following visit. Therefore, this study was using the VF results from the second visit of the subjects to determine the validity of SP in the estimation of normal global and pointwise threshold sensitivities. The difference in the time duration between both strategies was also determined in this study. The characteristics of threshold estimates were further investigated by evaluating their relationship with age and the spherical equivalent of the subjects. With regards to the determination of the threshold estimates which was associated with six functional regions described by Gonzalez de la Rosa et al. (2002a) for SPARK and eccentricity for SITA (Wild et al., 1999; Capris et al., 1999; Blumenthal et al., 2003), the effects of the functional regions and eccentricity of the test points on the differences of threshold estimates between the two strategies were also investigated.

2.1.3 Comparison between SPARK Precision and SITA Standard in Glaucoma Patients

The third experimental study in this thesis was aimed to investigate the extent of any differences of the global indices (i.e. MS, MD and PSD) and test duration between SP and SS in a group of glaucoma patients. A group of normal subjects was recruited as controls who were matched for age and refractive error with the glaucoma patients. Due to the fact that glaucoma is a common ocular disease exhibiting localized (i.e. focal) field loss (Sihota et al., 2007; Steele and Spry, 2009) which subsequently affects global indices of the threshold perimetry, patients with confirmed glaucoma diagnosis were recruited for this study. The pointwise differences of threshold estimates between the strategies were also evaluated in the glaucoma patients. The severity level of glaucoma can be different with different algorithms

(Bengtsson and Heijl, 1999; Budenz et al., 2002a; Heijl et al., 2000) which can affect the detection of the glaucomatous progression. The severity level of glaucoma using the Advanced Glaucoma Intervention Study (AGIS) severity scale (AGIS, 1994) and the size and depth of the glaucomatous field defect according to criteria recommended by HPA (Hodapp et al., 1993) were compared between the two strategies. The sensitivity and specificity of SP and SS in detecting glaucomatous VF defects were also determined and compared.

2.1.4 Comparison between SPARK Precision and SITA Standard in Subjects with Cataract

Cataract which is featured as the opacity of crystalline lens was shown to display diffuse VF loss (Rehman Siddqui et al., 2007; Koucheki et al., 2004; Ang et al., 2010). The final study of this thesis was to evaluate the influence of the cataract on the outcome of SP and SS in determining the global indices and test duration of the strategies in a group of patients with diagnosed cataract whereas another group of normal subjects was used a control. The global and pointwise threshold estimates were compared between SP and SS and the agreements were also determined for all the global indices.

2.2 Logistics

The research was conducted within the Faculty of Optometry and Vision Sciences, SEGi University, Petaling Jaya, Selangor, Malaysia. The studies had obtained approval from Aston University Research Ethics Committee (2nd of February 2016; #755). It is a fulfilment required for the author to be graduated as Doctor of Optometry from Aston University. The study was also given a research grant by SEGi University (SEGiIRF/2014-41/FOVS-2/51) for the expenses of the maintenance and services of the instrument and partly used for subjects' travelling expenses.

Faculty of Optometry and Vision Sciences in SEGi University was founded in 2010 and an Optometry Clinic was set up to serve as an optometry students training centre. Due to that fact

that it is located within a university campus with a mainly young population, patients with glaucoma were scarcely encountered. As the research needed to be completed within two years, extra effort was put in to recruit glaucoma patients. The ophthalmologists practising within 20 km of travelling distance as well as the local glaucoma society (Malaysia Glaucoma Society) were approached for the recruitment of glaucoma patients. The recruitment was also extended to the family members with glaucoma. Approximately one-third of glaucoma patients were recruited from the clinics of two ophthalmologists who had agreed to assist in the subject recruitment. The ophthalmologists are Dr Khor Sim Ee who practises in Thomson Hospital, adjacent to SEGi University and Associate Professor Dr Sushil Kumar Vasudevan from Faculty of Medicine, Universiti Teknologi MARA (UiTM), Sungai Buloh campus which is located about 12 km from SEGi University. All glaucoma patients recruited were patients with the confirmed diagnosis and were being treated by their ophthalmologist.

Subjects with cataract were recruited from university staff and patients attending the optometry clinic. All of the subjects were diagnosed following a routine optometric examination to confirm that they fulfilled the necessary inclusion and exclusion criteria. Normal controls were recruited after they had given their consent for the participation and followed with a standard eye examination confirming their eligibility.

As the VF testing required patients' concentration and longer visiting time as more than one hour was needed to complete the VF testing with both threshold strategies and an interval break, resistance of the patients to participate in this research mainly came from patients with cataract and elderly control subjects. Most of the cataract patients were unwilling to have another visit for the VF testing as they regarded the procedure was not necessary for the treatment of their cataractous eye(s). Also, there was a lack of patients with denser cataract as the treatment of the cataract is easily available and accessible in the urban area. Furthermore, it was a huge challenge to get a sufficient number of healthy elderly subjects who could be matched with the age of the cataract and glaucoma group. The prevalence of cataract

(Vashist et al., 2011; Song et al., 2018; Na et al., 2014) and glaucoma patients (Bourne et al., 2016; Rudnicka et al., 2006) increases with age. Thus, glaucoma and cataract patients recruited were mainly elderly patients and it was inevitably to have subjects who fulfilled the inclusion criteria but could not participate due to work commitments, travel difficulties or just unwillingness to participate. Some patients who are mainly the elderly normal subjects dropped out after the first visit. There was a period when the motherboard of the HFA model HFA II-745i which was used at the beginning of the study was faulty that the recruitment of the subjects had to come to a halt. A new model HFA III-830 was then used to replace the previous one but the delivery of the new instrument had taken more than a month.

With the dedication and enthusiasm of the author despite having to commit to a full-time job, more proactive approaches were taken to arrange after-hour clinical sessions at the convenience of the participants even during weekends or public holidays. In order to recruit more subjects, frequent visits to the ophthalmology clinics were made as well as conducting community vision screening. On top of these, a lot of patience was required while conducting the time-consuming VF testing on subjects. All these efforts paid off when the targeted number of glaucoma and cataract patients was achieved. Nevertheless, the most pleasant and biggest sense of achievement carrying out this research was the detection of some first diagnosed glaucoma patients during the process of the research.

CHAPTER 3

BETWEEN-VISIT AND WITHIN-STRATEGY THRESHOLD ESTIMATES ANALYSIS FOR SITA STANDARD AND SPARK PRECISION

3.1 Introduction

Threshold perimetry determines the minimum threshold to light at various locations of the VF (Jampel et al., 2011; Sample et al., 2011). It is a subjective psychophysical test which is dependent on the patient's co-operation and consistent response to achieve a reliable result. Even though the most part of standard automated perimetry has been computerized, it is still subject to patient's fatigue, inter-subject variability and high test-retest variability (Sharma et al., 2008; Nouri-Mahdavi et al., 2011; Johnson et al., 1988; Heijl and Drance, 1983; Gonzalez de la Rosa and Pareja, 1997; Hudson et al., 1994). Introduction of the fast threshold strategy in recent years has helped to produce more stable threshold estimates.

3.1.1 Variability

The variability of the VF test is affected by the depth of the VF defect (Heijl et al., 1989a; Boeglin et al., 1992; Chauhan and Johnson 1999; Henson et al., 2000; Blumenthal et al., 2003; Werner et al.,1989; Starita et al., 1987; Olsson et al., 1993; Fredette et al.,2015; Maddess, 2011; Artes et al., 2005; Gardiner et al., 2012), eccentricity of the test points (Heijl et al., 1989a; Boeglin et al., 1992; Chauhan and Johnson 1999; Wild et al., 1995; Blumenthal et al., 2003) and the test duration (Wild et al., 1999; Bengtsson and Heijl, 1998; Bengtsson and Heijl, 1999a; Aoki et al., 2007, Artes et al., 2002). Katz and Sommer (1987) found that variability is affected by age but Heijl et al. (1988) showed otherwise. Larger variability was exhibited in glaucomatous eyes than normal eyes (Chauhan et al., 1993; Flammer et al., 1984a) especially in the more advanced stage of glaucoma (Artes et al., 2005; Blumenthal et al., 2000; Gardiner et al., 2012). Variability of the threshold values was also found to be higher in patients with optic neuritis, ocular hypertension and also normal subjects with lower sensitivity (Henson et

al., 2000; Wild et al., 1999). The longer the test duration, the larger the variability of the test (Bengtsson et al., 1998; Aoki et al., 2007; Wild et al., 1999) which is mainly due to increased fatigue (Heijl and Drance, 1983; Hudson et al., 1994). The variability also could be associated with a reduced number of ganglion cells (Curcio and Allen, 1990; Chauhan and Johnson, 1999; Henson et al., 2000) or nerve fibres (Flammers et al., 1984a). It is also possibly related to poor fixation control (Henson and Bryson, 1991; Henson et al., 1996). Variability affects the ability of a test to detect the true change of VF from pathophysiologic fluctuation. Thus, it is clinically important to use the test that produces results with lower variability to increase the diagnostic sensitivity of detecting the true VF change. Introduction of fast threshold strategies such as Swedish Interactive Threshold Algorithms (SITA) or SPARK helped to shorten the testing time whilst maintaining high sensitivity and specificity (Garg, 2008; Budenz et al., 2002; Gonzalez de la Rosa et al., 2013). The fast threshold strategy could reduce the intra- and inter-test variability (Bengtsson and Heijl, 1999a). Nevertheless, the consistency of the short-duration VF test is affected if the patient lacks perimetric-experience (Pierre-Filho et al., 2006; Schimiti et al., 2002).

3.1.2 Learning Effect

It has been shown that perimetric results can be improved with practice, training and learning (Gloor et al., 1980; Werner et al., 1990). The “learning effect” in perimetric tests is manifested in improved threshold estimates, decreased variability of the threshold values and shorter testing time. It could be either a physiological process in the visual system or a psychological effect of the patients in decision-making (Saigal, 2011). The learning effect is most pronounced after the first 2 or 3 sessions of VF testing (Wood et al., 1987; Heijl et al., 1989; Searle et al., 1991; Heijl and Bengtsson 1996) and can even be shown after several years (Gardiner et al., 2008). Numerous studies have shown the existence of the learning effect using standard automated perimetry either in normal subjects (Wood et al., 1987; Heijl et al., 1989; Searle et al., 1991; Autzen and Work, 1990), patients with ocular hypertension (Wild et al., 1989; Werner et al., 1990) and glaucoma patients (Gloor et al., 1981; Werner et al., 1988; Kulze et al., 1990;

Wild et al., 1991; Heijl and Bengtsson, 1996). It still persists in SAP using fast threshold strategy (Castro et al., 2008; Aydin et al., 2015) even though both the test duration and variability are significantly reduced compared to Full Threshold strategy. Nevertheless, the learning effect was more prominent in Full Threshold compared to SS according to Yenice and Temel (2005). Non-conventional perimetry such as frequency doubling technology (FDT) perimetry (Fujimoto et al., 2002; Joson et al., 2002; Horani et al., 2002; Teresa et al., 2007; Heeg et al., 2003), flicker perimetry (Lamparter et al., 2011; Bernardi et al., 2007) and short-wavelength automated perimetry (SWAP) (Wild et al., 2006; Rossetti et al., 2006) have also been shown to exhibit a learning effect even though the patients tested all had prior experience in conventional SAP (Fujimoto et al.,2002; Zhong et al., 2008) or manual (Goldmann) perimetry (Werner et al., 1988). The perimetric-experience does not transfer to another perimetry test which uses different stimulus (Fredette et al., 2015) instead it was shown to be transferable between eyes (Searle et al., 1991; Heijl et al., 1989; Heijl and Bengtsson, 1996).

Global indices such as mean sensitivity (MS) or mean deviation (MD) can be improved to about 1 to 2 dB through learning especially between the first and second test in normal subjects (Wood et al., 1987; Heijl et al., 1989; Searle et al., 1991; Autzen and Work, 1990). The effect increases with eccentricity (Wood et al., 1987; Heijl et al., 1989; Searle et al., 1991; Wild et al., 1989; Werner et al., 1990; Heijl and Bengtsson, 1996) and the superior field appears to be affected more than the inferior field (Wood et al., 1987; Heijl et al., 1989) which may be due to patients having learned to be more aware to raise their eyelids during the perimetric test (Wood et al., 1987). The learning effect is also greater in the area where the VF defect is more severe (Kulze et al., 1990; Heijl and Bengtsson, 1996). This is also true for patients with myopia compared to emmetropes (Marra and Flammer, 1991). Aydin et al. (2015) found that the learning effect was associated with age and education level but not gender whereas Castro et al. (2008) disagreed with none of the afore factors linked with the learning effect in their sample.

3.1.3 Reliability Indices in Humphrey Field Analyser and Oculus Twinfield

Reliability is a paramount issue in standard automated perimetry (SAP) especially when it is used in monitoring the glaucomatous field changes. A perimetric test with low reliability affects the sensitivity and specificity of the test which could delay the medical treatment provided. Humphrey Field Analyzer (HFA) (Carl Zeiss Meditec, Dublin, CA, USA) in which SITA Standard (SS) is incorporated has several features to determine the reliability of the test. There are three classical indices used, i.e. fixation loss (FL) rate, false positive (FP) rate and false negative (FN) rate (Anderson and Patella, 1999). HFA uses Heijl-Krakau method (Heijl and Krakau, 1975; Heijl et al., 2012) determine FL rate where a stimulus is projected on to the predetermined blind spot. If the subject responded to the stimulus, a FL will be recorded. Classical FP error is when the subject responded without stimulus is projected. In SITA, with the intention to reduce the test duration, no catch trial is performed. Instead, the FP rate is determined by recording the subject's responses within the non-allowable time according to the subject's response time measured during the test. The non-allowable time starts immediately after the stimulus is presented and lasts for 180ms which is the minimum response time (Olsson et al., 1997) and it continues immediately after the response window until 180ms after the next stimulus presented (Olsson et al., 1997; Newkirk et al., 2006). High FP rate has been always associated with "trigger-happy" subjects. False negative error is determined through catch trials in which a suprathreshold stimulus of up to 20 dB higher than the threshold is shown at the location at the subject responded before (Bengtsson and Heijl, 2000). High FN indicates fatigue or lack of attention during the test.

Oculus Twinfield (Oculus, Wetzlar, Germany) used similar reliability indices except for FN which was ditched to reduce test duration. The FN rate provides the least information about patient attentiveness/cooperation among the three reliability indices but is associated more to the severity of glaucomatous field defects (Katz and Sommer, 1988; 1990; Katz et al., 1991a; Bengtsson and Heijl, 2000). It is therefore no longer a recommended criterion used to determine the reliability of the VF test using SITA according to the manufacturer's guideline

(Heijl et at., 2012; Carl Zeiss Meditec, 2010). The FL rate is determined using the catch trials but instead of blind spot, the central threshold is used. A stimulus which is brighter than the central threshold is projected onto the centre of the VF (4 red lights arranged in small diamond shape is used as fixation target). If the subject failed to respond to the light, a FL is recorded. The FP rate in Oculus perimeter is determined according to the classic way which only sound is produced but no stimulus is shown.

3.2 Objectives

The Swedish Interactive Threshold Algorithm (SITA) is a new generation strategy developed in the mid-1990s (Bengtsson et al., 1997). It is incorporated in the Humphrey Field Analyzer (HFA) (Carl Zeiss Meditec, Dublin, CA) and regarded universally as the gold standard for VF testing. SPARK, on the other hand, is a new generation threshold strategy developed for the Oculus perimeters (Oculus, Wetzlar, Germany) which was designed with the aim of reducing fluctuations in SAP.

The main objective for chapter 3 of this thesis was first to evaluate the reliability indices and between-visit variability of the threshold estimates of SITA Standard (SS) and SPARK Precision (SP) within each subject group i.e. glaucoma, cataract and healthy control subjects. Secondly, the influence of age on the between-visit within-strategy sensitivity variation was also determined.

To be specific, the aforementioned objectives could be achieved as followed:

a) To determine the number of eyes with an unreliable result within each subject group for SS and SP respectively considering the effects of the testing order and perimetric experience of the subjects.
b) To compare the MS from the 66 matching test points between the two visits within the same strategy.

c) To determine and compare the changes of the MS between-visit in perimetric-experienced and perimetric-inexperienced subjects.
d) To determine the pointwise sensitivity variation between-visit for the 66 test points in healthy control subjects.
e) To find out the correlation between age and the sensitivity changes between-visit for both SS and SP in healthy control subjects.

3.3 Methods

3.3.1 Research Participants

This was a prospective, cross-sectional study and the recruitment of subjects adhered to the tenants of the Declaration of Helsinki and approved by Aston University Research Ethics Committee (2nd of February 2016; #755). This study included 3 groups of participants: glaucoma and suspected glaucoma patients, cataract patients and 2 groups of normal subjects as controls (two control groups were required as glaucoma patients are typically younger than cataract patients). All participants were recruited from eye clinics, hospitals, optometry clinics and the local glaucoma society (Kuala Lumpur, Malaysia). The glaucoma group consisted of patients with stable open or closed angle glaucoma. All of the glaucoma patients had typical glaucomatous cupping of the optic disc and/or VF loss on 30-2 or 24-2 HFA SS testing and were medicated with a confirmed diagnosis by their ophthalmologist. Suspected glaucoma patients had a family history of glaucoma and/or suspicious discs but no significant structural changes and normal IOP and VFs and not medicated at the time of recruitment. Cataract patients had significant opacities of the crystalline lens [at least nuclear opalescence (NO) / nuclear colour (NC) grade 3 or posterior subcapsular cataract (P) grade 2 with Lens Opacities Classification System (LOCS) III (Chylack et al., 1993; Karbassi et al., 1993)] with healthy fundus and best corrected visual acuity (BCVA) worse than 6/6 and IOP below 21mmHg. Normal subjects showed no signs of glaucoma with full VFs, IOP ≤ 21mmHg with non-contact tonometry, normal optic discs (i.e., no localized rim loss, optic disc haemorrhage, cup/disc ratio

≤ 0.6 or cup/disc asymmetry ≤ 0.2, notches, localized pallor, or nerve fibre layer defects), and exhibited no family history of glaucoma.

All subjects had given their written informed consent before they enrolled in this study. The participants of the control group underwent a comprehensive eye examination to determine their eligibility for this study after their consent. Cataract patients were recruited once they were diagnosed during a routine optometric eye examination. Other inclusion criteria are listed below:

a) Minimum and maximum age for participation 18 to 80 years.
b) Best corrected visual acuity (BCVA) 6/12 or better for controls and glaucoma patients (except for cataract patients: 6/6 > BCVA ≥ 6/18).
c) Refractive errors below +/-6DS in sphere, 2.5DC or less in astigmatism (3.5DC or less for cataract subjects).
d) No history of intraocular surgery complications and no ocular disease (except glaucoma or cataract) and systemic illness that could affect visual fields (e.g. pituitary lesions, demyelinating diseases, HIV+, AIDS, or diabetic retinopathy)

Subjects were excluded if they were diagnosed with uncontrolled diabetes mellitus, untreated hypertension, or other systemic disease, pregnant or nursing, use of drugs potentially affecting reaction time, use of alcohol, nicotine or caffeine less than 2 hours before their perimetric examination, taking medications that known to affect VA and low reliability indices from their VF tests (FP rate < 33%, FN rate < 20% and FL < 20%). Normal subjects were excluded if any significant VF defect was detected in two consecutive visits using HFA.

3.3.2 Methods and Procedures

All the normal controls in this study underwent a standard routine eye examination which included VA test, refraction, non-contact tonometry with Nidek NCT-510 (Nidek, AiChi, Japan),

corneal pachymetry and anterior chamber angle imaging using Scheimpflug topographer (TMS-5, Tomey, Nagoya, Japan), external and internal eye examinations using slit lamp biomicroscopy and non-mydriatic retinal photography with a fundus camera (Topcon TRC-NW300, Topcon, Tokyo, Japan) to confirm their eligibility for the study after they had given their written consent.

After confirmation of their eligibility to this study, all the subjects including glaucoma and cataract patients underwent VF assessments with the Oculus Twinfield 2 (Oculus, Wetzlar, Germany) using SPARK Precision (SP) and the Humphrey Field Analyzer (HFA) (Carl Zeiss Meditec, Dublin, CA) using SITA Standard (SS) within a day. The 30-2 test point pattern was used in SITA with a total of 76 test points covering the central 30-degree field with a square grid of 6-degree separation. SPARK was using an almost similar grid which has a total of 66 test points (30° x 24°) and as compared to 30-2, it is missing the uppermost and bottommost rows and two points located adjacent to blind spots (Gonzalez de la Rosa et al., 2013). A break of about 10 minutes was given between the two tests. The right eye or better eye was chosen to undergo the VF test first for each subject if both eyes were eligible and consent was given for both eyes to be tested. All the assessments were conducted by a qualified optometrist. The order of SITA and SPARK testing was randomised among subjects to minimise the learning effects but remained identical between the visits of the same participant. The data from both visits were collected and analysed.

3.3.2.1 Visual field assessment with SITA Standard

The HFA model HFA II-745i or HFA III-830 (Figure 3.1) was used which has a bowl shape of 30cm testing distance. Background illumination was recorded as 31.5asb (10cd/m^2) and a white Goldmann size III stimulus (4mm^2) was selected. Stimulus duration was 200ms and the stimulus interval was set to adaptive (0.6 to 2s). Fixation control in HFA was using the Heijl-Krakau method (Heijl and Krakau, 1975). A video eye monitor was used in HFA II-745i and

HFA III-830 to observe the patient during testing. A yellow central round light was used as the fixation target for the VF test.

At first, subjects were asked to sit comfortably in front of the perimeter. The contralateral eye was occluded. Each subject's near correction was calculated according to their distance correction and age using the calculation recommended by manufacturers with a testing distance of 30cm (Carl Zeiss, 2005; Oculus Twinfield 56920 instruction manual, Wetzlar-Dutenhofen, Germany). The trial lens was placed as close to the subject's eye as possible without touching the eyelashes. The chair and table heights were adjusted so the subject could comfortably put his/her chin firmly on the chin rest and the forehead against the forehead rest. The subject's eye position was aligned using the video eye monitor to place the cross in the centre of the pupil by adjusting the chin rest. Prior to starting the SS test, the subject was given a buzzer and the standard test instructions recommended by the manufacturer's operating manual were briefed in order to allow each subject to give a neutral/natural response (Kutzko et al., 2000). Subject's fixation was continually monitored throughout the test and he/she was reminded not to shift his/her position from the chin and forehead rest. After completing the first eye, a short break of about 2 minutes was given if the second eye was to be tested as well. The room lights were switched off for the entire VF testing.

Figure 3.1: A subject is undergoing a visual field test using SITA Standard, Humphrey Field Analyser HFA III-830 (the room light was switched on for the purpose of photo taking).

3.3.2.2 Visual field assessment with SPARK Precision

The Oculus Twinfield 2 perimeter (Figure 3.2) was used for assessing the VF by SP. It is bowl-shaped with a testing distance of 30 cm but the background illumination is set to 31.8 asb. Similar to SITA the stimulus size of a white Goldmann size III and 200 ms stimulus duration were used. Stimulus interval was recorded within the range of 1.6 sec. Four small LED lights arranged in diamond-shaped were used as the fixation target. Fixation of the eye was controlled using the central threshold and a video eye monitor. The room lights were switched off during the examination.

As the routine of a VF test, the subject was sitting comfortably in front of the perimeter and occluder was used to cover the non-test eye. Lens holder was inserted in the holding device

provided near the chinrest if the subject needs the correction. Patient's current distance correction of glasses and age were taking into account for the calculation of the corrective lens. The same correction lens was used for both perimeters. The corrective lens was placed as close to the subject's eye as possible without touching the eyelashes by adjusting the position of the forehead rest. The table height and chin rest were adjusted to make sure the subject was comfortably seated during the test. The same instruction as recommended by HFA's manufacturer was used as well. Foveal threshold test was started as the first procedure before the SPARK strategy test was initiated. An approximately 2-minute break was given after the first eye was completed and if the other eye was also tested.

Figure 3.2: Visual field test using SPARK Precision, Oculus Twinfield (the room light was switched on for the purpose of photo taking).

3.3.2.3 Analysis of visual field results

The results that have low-reliability indices (FP rate > 33%, FN rate > 20% and FL > 20%) (Cubbidge, 2005; Rao et al., 2017) were identified as unreliable and if repeated unreliable

results were obtained in the second visit for any of the VF test, the subject would be excluded. The FN rate is not available in SPARK. Only false positive rate > 33% and FL > 20% were used for SPARK results. If both eyes fulfilled the inclusion and exclusion criteria and their results were acceptable in term of reliability, the better eye was chosen for a normal subject whereas the selected eye from glaucoma patients was biased toward milder VF defect if both eyes exhibited significant glaucomatous field defects. The chosen cataract eyes in this study were preferred towards the eye exhibiting the better BCVA if bilateral cataract was diagnosed and both eyes had monocular VA worse than 6/6. If there was no difference between both eyes, the second tested eye was selected for the study. Only one eye of each patient was used for the data analysis of this chapter.

If normal subject exhibited abnormal VF results in two consecutive visits either in SS or SP then they were excluded. The criteria used to determine abnormal VF in SS followed the criteria recommended by Hodapp-Parrish-Anderson (HPA) (Hodapp et al., 1993). A subject was considered having abnormal VFs if either one of the following criteria was found in their SS results:

a) A cluster of three or more non-edge points in either hemifield on pattern deviation probability map with sensitivity found in <5% of the normal population (p<5%) with at least one of the defective points having p<1%,
b) Pattern standard deviation (PSD) had a value found in <5% of the normal population (p<5%),
c) Glaucoma hemifield test (GHT) showed “outside normal limits”.

Whereas abnormal VF with SP was determined based on the criteria recommended by Gonzalez de la Rosa et al. (2013) in which at least 95% specificity was achieved

a) Mean deviation (MD) with a value lower than -2.3 dB,
b) Pattern standard deviation (PSD) with a value of more than 1.8 dB,
c) More than 5 abnormal total deviation points with deviation worse than 5 dB.

(*Obtained through written communication from Gonzalez de la Rosa, November 30, 2013)

The sensitivities of the matching 66 test points in SS corresponding to the test points used in SP were used to calculate the MS of each test. The uppermost and bottommost rows of test points and the two test points located at the blind spot in SITA 30-2 were excluded in the analysis. The central threshold in SPARK was also excluded in the calculation of MS. The left eye results were converted into right eye format. The MS for the 66 test points from each strategy test was recorded and compared between the visits within each subject group.

3.3.3 Statistical Analysis

The total number of normal subjects required in this study was 40 per group which is based on the calculation using GPower 3.1.3 (Faul et al., 2007; Prajapati et al., 2010) with an alpha value of 0.05 and statistical power of 80%. There were initially two control groups targeted: subjects below 40 years old and subjects aged 40 years old and above. The effect size was calculated using the mean and standard deviation of MD and testing time from Schimiti et al. (2002). Whereas the total number of glaucoma subjects was estimated using data obtained from Sekhar et al. (2000) with an alpha value of 0.05 and achieved a statistical power of 80% and at least 30 patients need to be recruited for glaucoma group with a forecast of 10% drop out rate. The similar targeted number of subjects (30) was used for cataract group as the required data for the estimation of minimum sample size in this group cannot be found from the previous studies up to June 2018.

All data were analysed using SPSS version 22 (IBM Corp, Armonk, NY) and normality of the data was determined first by using the Shapiro-Wilk test. Mean and standard deviations were used to describe normally distributed continuous variables whereas median and range were used for non-normally distributed continuous variables. Chi-square test was used to compare categorical data and Cramer's V to determine the association between the categorical data. Paired t-test or Wilcoxon Signed Rank test was used for comparison of continuous variables between-visit within-strategy including within groups according to perimetric-experience.

Means and standard deviations of the mean difference of the MS were used to determine the pointwise variation between the visits within the same strategy for the control group. The pointwise variation of MS was compared according to the location and eccentricity of the test points. Spearman correlation coefficient was used to determine the association between the non-normally distributed continuous variables including the relationship between the age and between-visit threshold variation.

3.4 Results

A total of 158 subjects participated in this study. Of these, 44 were glaucoma patients or glaucoma suspect, 31 had cataract and a total of 83 healthy controls. They had all completed at least 2 visits of VF tests.

3.4.1. Glaucoma Group

There was a total of 44 subjects recruited into glaucoma group which comprised of 34 open-angle glaucomas (OAG), five angle closure glaucoma (ACG) and five glaucoma suspects. All the 39 glaucoma patients (OAG and ACG) were either under anti-glaucoma medication or had undergone glaucoma treatment surgery. The other five subjects were classified as glaucoma suspects who exhibited suspicious discs and/or have a family history of glaucoma but had no treatment and/or medication prescribed by ophthalmologists and they did not show significant VF defects. Their IOPs were also below 21mmHg. Among all the 44 subjects, six subjects (five OAG and one ACG) had the monofocal intraocular lens. Among the rest of the 38 phakic subjects, eight were found to have cataract (including one glaucoma suspect) with two of them were graded NO/NC grade 4 (LOCS III) and the remainders were NO/NC grade 3 or less including a glaucoma patient who had combined nuclear and cortical cataract [NO/NC grade 3 and cortical (C) grade 2]. One OAG patient had trabeculectomy and three out of five ACG patients had undergone iridectomy. All the surgeries were uneventful and completed for more than 6 months ago. Out of 34 OAG patients, 30 OAG patients and three out of five ACG had exhibited significant VF defects using SS according to HPA criteria (Hodapp et al., 1993).

Twenty-one of them were female and 23 were male. Among the glaucoma suspects, two of them were females and three were males. The BCVA of all glaucoma patients and suspects ranged from 6/4 to 6/12 (median = 6/6).

Twelve of the OAG glaucoma patients were initially diagnosed in the optometry clinic when they presented for this research and hence were subsequently recruited to participate in it. All of them were referred to ophthalmologists and had their diagnosis confirmed. The remaining 27 glaucoma patients were existing patients from the following centres:

a) Glaucoma specialist clinic under Associate Professor Dr Sushil Kumar Vasudevan from Faculty of Medicine, Universiti Teknologi MARA (UiTM), Sungai Buloh campus (7 patients),
b) Adjacent medical centre under a general ophthalmologist (Dr Khor Sim Ee in Thomson Hospital Kota Damansara, Petaling Jaya) (8 patients),
c) Government hospitals or other private ophthalmology clinics (12 patients) where the patients were first diagnosed but subsequently recruited when presented in the optometry clinic for a routine eye examination.

Four out of five glaucoma suspects were recruited from Dr Khor's clinic and another one was a walk-in patient from the optometry clinic. All these five glaucoma suspects had shown no significant VF defect with either SS or SP.

3.4.2 Cataract Group

All cataract subjects were recruited in an optometry clinic with a total of 31 subjects (19 females and 12 males) showing significant crystalline lens opacities. All of them had reliable results for both VF tests at least in their second visit. Among them, 25 were diagnosed with nuclear cataract (NO/NC grade 3 and above), three with nuclear and cortical cataract (C grade 2) and another three had posterior subcapsular cataract (P grade 2) (LOCS III). Their BCVA ranged from 6/7.5 to 6/18 (Median = 6/9). Out of 31, 20 of them (64.5%) had their BCVA at least 6/9.

Twelve of them were recommended to undergo cataract removal surgery based on the severity of their cataract, visual demand, age and occupational needs. By the date of the completion of this research, five of them had undergone cataract removal surgery. The rest of the seven subjects have been delaying the surgery due to financial issues or work commitments. The other 19 non-referral cataract subjects remained under periodical monitoring.

3.4.3 Control Subjects

Eighty-three normal subjects were recruited from university staff, students and patients attending the optometry clinic. All of them had completed at least two visits of VF tests with no consecutive VF defect shown in SS and SP. All control subjects had achieved at least 6/6 in BCVA. It comprised of 49 females and 34 males. Among them, 42 subjects were below 40 years old and 41 subjects were 40 years old and above.

3.4.4 Between-visit Within-strategy Within-group Analysis

3.4.4.1 Glaucoma group

Eleven out of 44 glaucoma patients did not have reliable VF test results in their first sitting. Of these, six glaucoma patients (13.6%) had unreliable results with SS (three had FL>20%, two had FN>20% and one had both FL>20% and FN>20%) and three (6.8%) with SP with two of them had FL>20% and another subject had FP>33%. The remaining two subjects (4.5%) had unreliable results in both SS and SP (one had FL and FN > 20% and another one had FL>20% and FP>33% in SS whereas their SP results showed FL>20%). No statistically significant difference in the number of subjects with unreliable results between SS and SP (Chi-square test: $p = 0.179$). By comparing the results with FL > 20%, six subjects with SS and four subjects with SP were found. Each strategy had one subject with FP > 33%. There were four subjects who had FN > 20% using SS.

Only nine glaucoma patients (20.5%) had no experience in VF test prior to their first visit and five out of the nine had unreliable VF tests in their first visit (see aforementioned details). The

reliability of the SITA results was found to be statistically correlated to the prior experience of the subjects in performing VF test (Cramer's V: $p = 0.022$) but not for the reliability of SPARK results (Cramer's V: $p = 0.250$).

There were 20 glaucoma subjects who started with SP in their order of the tests whereas another 24 subjects started with SS. Among the 11 glaucoma subjects that had unreliable results in the first visit, six started with SS and five started with SP. The reliability of the SS and SP results was not statistically significant correlated to the order of the test (Cramer's V: $p = 0.775$ for SS; $p = 0.225$ for SP). Out of the five perimetric-inexperienced subjects prior to their first visit who also obtained unreliable results in their first visit, four of which had started with SS but two had unreliable results in SS, one had unreliable results in both SS and SP and another one was only with SP. Another subject had unreliable result in SS but he started his first VF test with SP.

For the analysis between-visit in the glaucoma group only data from n=33 glaucoma patients who had reliable results using SS and SP in both two visits (FL < 20%, FP < 33% and FN < 20%). Among them, 25 subjects were diagnosed as OAG, four subjects were ACG and another four subjects were glaucoma suspects. Their mean age was 52.8 years (SD 11.4 years; range 30 to 70 years). The mean duration between the two visits was 10.6 days (SD 7.6 days; range 2 to 35 days). Only four of them (12.1%) had no experience of perimetry prior to the first visit with two apiece started either with SS or SP. Among these 33 glaucoma patients, those started the order of the test with SS were 18 subjects and those with SP were 15 subjects.

Shapiro-Wilk test of the MS of both test strategies showed they were not normally distributed ($p < 0.01$). The median and ranges of MS of the two threshold strategies are shown in Table 3.1. Wilcoxon Signed Rank test was used for the comparison of MS between both visits within the same strategy.

Table 3.1: Median and range of MS of SITA Standard and SPARK Precision in each visit of glaucoma group

	MS SITA (dB)			MS SPARK (dB)		
	Median	Range		Median	Range	
		Min	Max		Min	Max
Visit 1	27.79	13.97	31.67	28.67	13.73	32.89
Visit 2	28.42	11.24	31.45	28.68	11.55	33.09

The MS of the second visit using SS were significantly better than the first visit (Wilcoxon Signed Rank test: $p = 0.001$) but no significant difference was found between visits of the MS using SP (Wilcoxon Signed Rank test: $p = 0.379$).

3.4.4.2 Cataract group

Among the 31 subjects with cataract, six subjects (19.4%) had unreliable VF results in their first visit. Out of the six, four of them (12.9%) had unreliable results with SS (three subjects had FL > 20% and one had FN > 20%) and two (6.5%) with SP (one had both FL > 20% and FP > 33% and another one had only FL > 20%). No statistically significant difference was found between SS and SP in regards to the number of subjects with unreliable result in their first visit (Chi-square test: $p = 0.574$).

Only five out of 31 cataract subjects had prior experience in VF testing and all five of them had reliable VF results in both visits. Out of the 26 inexperienced subjects, six subjects had unreliable VF results in their first visit. No statistically significant relationship was found between the reliability of the VF test and experience of the subjects in performing VF test (Cramer's V: $p = 0.347$ for SS, $p = 0.521$ for SP).

Sixteen cataract subjects had started the VF test with SP and another 15 subjects started with SS. Among the six cataract subjects with unreliable results in their first visit, three subjects

started with SS whereas another 3 started with SP. No statistically significant correlation between the reliability of the VF test and the testing order (Cramer's V: p = 0.945 for SS, p = 0.962 for SP).

A total of 25 cataract subjects who had reliable VF results in both visits was used in the analysis of mean sensitivities between-visit within-strategy. Of these, 19 had nuclear cataracts, three had a combination of cortical and nuclear cataract and another three had posterior subcapsular cataract. Their mean age was 61.1 years (SD 10.2 years; range 31 to 76 years). The mean of the interval period between the two visits was 8.3 days (SD 4.2 days; range 1 to 22 days). There were 20 subjects (80%) who had not performed any VF test prior to the first visit and the rest of five subjects were experienced subjects. Twelve subjects started with SS and another 13 subjects started with SP. Among the subjects without perimetric experience, nine started with SS and 11 started with SP.

Shapiro-Wilk test showed mean sensitivities of SS were normally distributed (p = 0.172 for first visit; p = 0.231 for second visit) but mean sensitivities of SP were not normally distributed (p = 0.013 for first visit). Paired t-test was used for the comparison within SS and Wilcoxon Signed Rank test was used for SP comparison. The means and SDs of the MS of SS and medians and ranges of the MS of SP are shown in Table 3.2 and 3.3 respectively.

Table 3.2: Comparison of mean sensitivity of SITA Standard between visits in cataract group

	MS of SITA Standard (dB)	
	Mean	SD
Visit 1	26.41	2.19
Visit 2	27.36	1.72

Table 3.3: Comparison of mean sensitivity of SPARK Precision between visits in cataract group

	MS of SPARK Precision (dB)		
	Median	Range	
		Min	Max
Visit 1	27.30	19.73	30.86
Visit 2	28.05	23.41	30.97

Both SS (paired t-test: t = -2.361, df = 24, p = 0.027) and SP (Wilcoxon Signed Rank test: p = 0.004) showed statistical significant improved MS in the second visit compared to the first visits for the cataract group.

3.4.4.3 Control subjects

Of a total of 83 recruited control subjects, 14 subjects (16.9%) had to be excluded due to unreliable results in the first visit. Out of the 14, ten subjects (12.0%) had unreliable result with SS (FL > 20%), one subject (1.2%) with SP (FL > 20%) and another one subject had both unreliable results with SS (FL > 20%) and SP (FP > 33%). The remaining two subjects had shown artefact VF defects each with SS and SP despite obtained acceptable reliability indices. Even though statistically no significant difference was found in the number of subjects with unreliable results between SS and SP (Chi-square test: p = 0.344), there was a clinically significant difference of the number of eyes with FL>20% between SS and SP (11 in SITA vs one in SPARK). There was only one subject had FP>33% with SP but none with SS.

Thirty subjects (36.1%) were identified as the first timers in VF test but only six had unreliable results with SS and one with SP in their first visit. Prior experience in VF test had no statistically significant correlation to the reliability of the VF test (Cramer's V: p = 0.280 for SS; p = 0.918 for SP).

Forty-three subjects had started the VF test with SS and the remaining 40 started with SP. Among the 30 perimetric-inexperienced subjects, the number of subjects started with SS was equal to the number of subjects started with SP. No statistically significant correlation was found between the reliability of the VF test and the order of the tests (Cramer's V: $p = 0.625$ for SS, $p = 0.600$ for SP).

After excluding the subjects with unreliable results in the first visit, the total number of control subjects used for the analysis between-visit is 69. Their mean age was 37.6 years (SD 16.0 years; Range 20 to 71 years). Twenty-three of them (33.3%) had no prior perimetric experience. The testing order was approximately equally divided among the subjects with 34 subjects started with SS and 35 started with SP.

The data collected for MS of SS exhibited normal data distribution ($p = 0.265$ for first visit; $p = 0.831$ for second visit) but not for SP ($p = 0.005$ for first visit; $p = 0.002$ for second visit) according to Shapiro-Wilk test. The between-visit comparison of group mean and SD of MS of SS in the control group is shown in Table 3.4 whereas Table 3.5 shows the comparison of MS of SP.

Table 3.4: Comparison of mean sensitivity of SITA Standard between visits in control group

	MS of SITA Standard (dB)	
	Mean	SD
Visit 1	30.58	1.40
Visit 2	30.96	1.29

Table 3.5: Comparison of mean sensitivity using SPARK Precision between visits in control group

	MS of SPARK Precision (dB)		
	Median	Range	
		Min	Max
Visit 1	32.20	28.82	34.27
Visit 2	32.18	28.67	34.32

The MS of SS in the second visit had a statistically significant increase compared to the first visit (paired t-test: t = -3.038, df = 68, p = 0.003) but no statistically significant change was found between visits when using SP (Wilcoxon Signed Rank test: p = 0.522).

3.4.5 Comparison Within and Between Perimetric-experienced and Inexperienced Subjects

For this analysis, we pooled data from all the 127 participants with reliable results in both visits (33 glaucomas, 25 cataracts and 69 controls). Among them, 80 (63%) were perimetric-experienced subjects (29 glaucomas, 5 cataracts and 46 controls) and remaining (37%) were new to perimetry before the study. Comparison of the MS between the visits within all experienced subjects showed a statistically significant increase of MS of SS (Wilcoxon Signed Rank test: p < 0.001) but not for SP (Wilcoxon Signed Rank test: p = 0.301). Subsequently, comparison within the inexperienced subjects showed statistically significant improved MS in the second visit for both strategies (Wilcoxon Signed Rank test: p < 0.001 for SS, p = 0.027 for SP) (Table 3.6).

Comparisons of within-visit between the MS of experienced and inexperienced groups showed experienced subjects had statistically higher MS than inexperienced subjects with either SS or SP (Mann-Whitney: p = 0.006 for SS in visit 1; p = 0.002 for SP in both visits) except for the

second visit of SS, no significant difference was shown between the experienced and inexperienced group (Mann-Whitney: p = 0.072).

Table 3.6: Comparison of MS between visits according to the perimetric experience of subjects in all three groups (n = 127)

Strategy test		Mean sensitivity (dB)						Mann Whitney test
		Experience			Inexperience			
		Median	Range		Median	Range		p
			Min	Max		Min	Max	
SITA Standard	Visit 1	30.08	13.97	33.06	28.42	22.02	32.98	***0.006***
	Visit 2	30.24	11.24	33.61	29.55	24.36	33.32	0.072
Wilcoxon Signed Rank test	p		***<0.001***			***<0.001***		
SPARK Precision	Visit 1	31.48	13.73	34.27	29.42	19.73	33.92	***0.002***
	Visit 2	31.44	11.55	34.32	29.85	22.18	33.48	***0.002***
Wilcoxon Signed Rank test	p		0.301			***0.027***		

* Bold and italic p-value indicates a significant difference

Further comparisons of between-visit within the perimetric-experienced control subjects only (n = 46) were conducted which showed a statistically significant improvement of MS of SS (paired t-test: t = -2.074, df = 45, p = 0.044) but not with SP (Wilcoxon Signed Rank test: p = 0.563) in second visit. Similar results shown in inexperienced control subjects (n = 23) that a statistically significant improvement of MS of SS (Wilcoxon Signed Rank test: p = 0.013) but not with SP (Wilcoxon Signed Rank test: p = 0.693) between the visits (Table 3.7).

The MS of each strategy was also compared between experience and inexperience control groups within the same visit. It showed the experience normal group had statistically significant higher MS than inexperience group in both visits either with SS or SP (Table 3.8).

Table 3.7: Comparison of MS between-visit within control subjects according to perimetric-experience (n = 69)

Strategy test		Mean sensitivity (dB)					
		Experience			Inexperience		
		Mean*/ Median	SD*/Range		Mean*/ Median	SD*/Range	
			Min	Max		Min	Max
SITA Standard	Visit 1	30.96*	1.14*		29.52	27.14	32.98
	Visit 2	31.19*	1.26*		30.50*/ 30.47	1.23* 27.32	33.32
Paired t test^/Wilcoxon Signed Rank test	p	***0.044***^				***0.013***	
SPARK Precision	Visit 1	32.60	29.30	34.27	30.27	28.82	33.92
	Visit 2	32.81	29.80	34.32	30.55	28.67	33.48
Wilcoxon Signed Rank test	p	0.563				0.693	

* Bold and italic p-value indicates a significant difference

Table 3.8: Statistical comparison within-visit between perimetric-experience and perimetric-inexperience group in control subjects

Visit	Strategy test	t*/Z	df	p
Visit 1	SS	-3.183	-	***0.001***
	SP	-4.086	-	***<0.001***
Visit 2	SS	2.160*	67	***0.034^***
	SP	-4.405	-	***<0.001***

^ Unpaired t-test
The rest used Mann Whitney test
* Bold and italic p value indicates a significant difference

3.4.6 Pointwise Analysis Between-visit Within-strategy in Control Subjects

Further analysis was conducted on all 66 testing points for the between-visit MS changes of each test strategy within the control group. To do so, the mean difference between visits was calculated as follows:

Mean difference of MS = MS (visit 2) – MS (visit 1) (Eq. 3.1)

Results and SDs of the differences between the two visits in each test point are detailed in Figure 3.3 for SS and Figure 3.4 for SP.

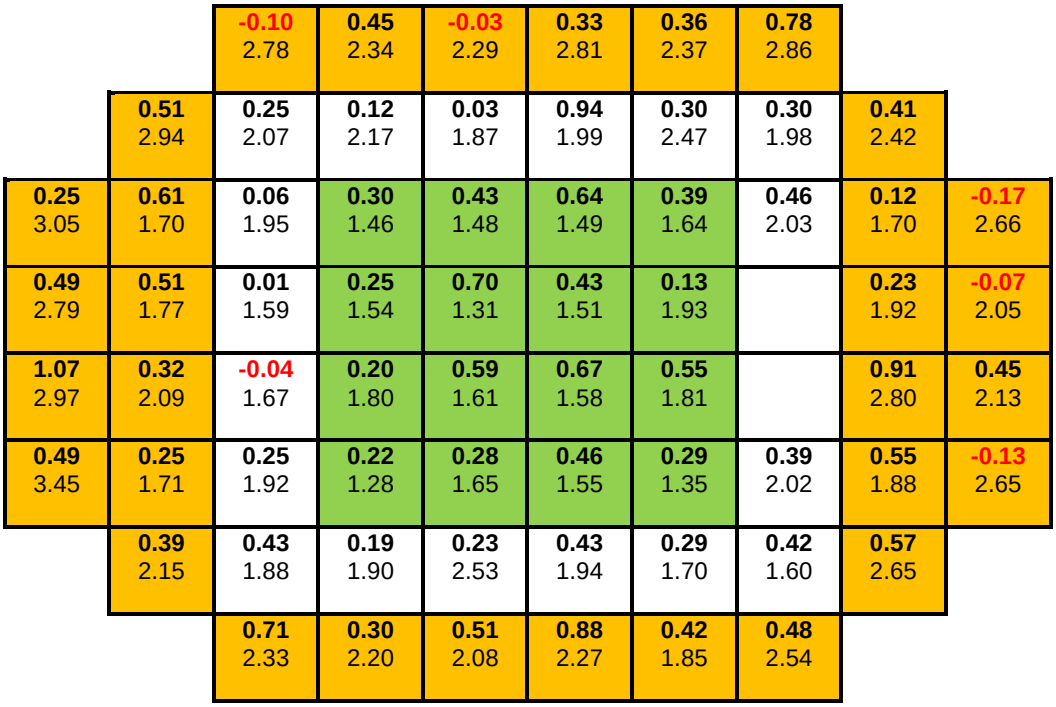


Figure 3.3: Mean difference and standard deviation of the difference of MS between the two visits at the 66 test points of SS which are matched to SP test points (as the field view of right eye) in normal subjects

* Bolded numbers – Mean difference of MS between the visits for SS; unbolded numbers – standard deviation of the difference of MS between visits
* Red numbers indicate higher MS in the first visit compared to the second visit
* The colour represents the groups of the test points according to the location of test points in central < 10° (green), mid-peripheral 10° - 20° (white) and peripheral > 20° (orange)

Figure 3.4: Mean difference and standard deviation of the difference of MS between the two visits at the 66 test points of SP (as the field view of right eye) in normal subjects

* Bold numbers – Mean difference of MS between the two visits for SP; unbolded numbers – standard deviation of the difference of MS between the two visits for SP
* Red numbers indicate higher MS in the first visit compared to the second visit
* The colours represent the functional map of VF defined by Gonzalez de la Rosa et al. (2002)

The majority of test points showed higher MS on the second visit when using SS except for five peripherals and one mid-peripheral test location. The SDs appear to be larger for the peripheral test locations compared to central ones. A comparison of the between-visit MS variation among the three areas (central < 10°, mid-peripheral 10 - 20° and peripheral > 20°) using SS showed no statistically significant difference (Kruskal Wallis test: Chi-square = 4.767, p = 0.092) (Table 3.9).

Table 3.9: Comparison of between-visit mean difference of mean sensitivity according to the location using SITA Standard

	Mean difference of MS between visits using SS (dB)	
	Mean	SE
Central	0.41	0.05
Mid-peripheral	0.28	0.05
Peripheral	0.40	0.05

Between-visit pointwise variations for SP were analyzed and it was found that more test points showed a reduction of MS in the second visit especially test points located in the superior VF. The pointwise SD for each test point was on average narrower compared to those obtained for SS. There was also no prominent larger SD in peripheral test points compared to central points. Comparison according to central, mid-peripheral and peripheral points was also conducted (Table 3.10). The data of the mean differences were normally distributed according to the Shapiro-Wilk test ($p > 0.05$) (Appendix A3.1). Hence, the comparison was conducted by a one-way ANOVA which showed a statistically significant difference between the mean differences of central, mid-peripheral and peripheral VF areas (one-way ANOVA: $F = 3.768$, $df = 2$, $p = 0.028$). Post hoc analysis (Tukey HSD) revealed the MS of the centrally located test points improved more than peripheral test points ($p = 0.03$) in their second visit.

Table 3.10: Comparison of between-visit mean difference of mean sensitivity according to location using SPARK Precision

	Mean difference of MS between visits using SPARK Precision (dB)	
	Mean	SE
Central	0.11	0.04
Mid-peripheral	0.07	0.03
Peripheral	0.00	0.02

The changes in the MS between visits were also analyzed according to the functional map defined by Gonzalez de la Rosa et al. (2002). Comparison among the six zones (superior, superior nasal, inferior nasal, inferior, temporal and central) showed statistically significant difference with one-way ANOVA ($F = 3.770$, $df = 5$, $p = 0.005$) (Table 3.11). The test points located in the superior zone showed poorer MS in their second visit which was significant different compared to superior nasal and temporal test points (Post-hoc analysis Tukey HSD: $p = 0.002$ for superior-superior nasal, $p = 0.025$ for superior-temporal).

Table 3.11: Mean differences of MS between visits according to functional map of VF defined by Gonzalez de la Rosa et al. (2002)

	Mean (dB)	SE (dB)
Superior	-0.11	0.03
Superior nasal	0.11	0.04
Inferior nasal	0.03	0.03
Inferior	0.04	0.04
Temporal	0.08	0.04
Central	0.11	0.04

3.4.7 Comparison of Between-visit Pointwise Variation between Strategies in Control Subjects

The comparison between the two strategies for the average of pointwise variations between-visit of all 66 test points showed SS had statistically significant larger variability (Paired t-test: $t = 9.175$, $df = 65$, $p < 0.001$) compared to using SP in normal subjects as shown in Table 3.12.

Table 3.12: Mean and standard deviation of pointwise mean differences between-visit using SS and SP

	Type of strategy	Mean	SE
Pointwise mean difference (dB)	SITA	0.37	0.03
	SPARK	0.04	0.02

3.4.8 Relationship between Age and Variation of Mean Sensitivity Between-visit in Control Subjects

Further analysis of the relationship between the variability between-visit with age in control subjects was conducted. It showed the MS variations between-visit within control subjects were not statistically significant correlated with age neither for SS nor SP (Spearman correlation coefficient: $rho = 0.106$, $p = 0.386$ for SS; $rho = 0.063$, $p = 0.607$ for SP). The medians and ranges of the MS variation between-visit in control subjects are given in Table 3.13.

Table 3.13: The median and ranges of pointwise variation of MS between visits in control group (n = 69)

	Pointwise variation between-visit (dB)			Age (yrs)		
	Median	Range		Median	Range	
		Min	Max		Min	Max
SS	0.32	-3.18	3.71	30.0	20.0	71.0
SP	0.12	-2.62	1.97			

3.5 Discussion

In this study, the consistency between-visit for each strategy, SPARK Precision (SP) and SITA Standard (SS) was evaluated to justify the necessity of the familiarization process for the VF test before the comparison study between the two fast threshold strategies to be carried out in the following chapters. The evaluation was performed by determining the significance of the MS changes between the two visits.

By excluding subjects with unreliable results in either of the two visits of this study, the improvement of the MS between the two visits was found to be statistically significant in SS for all three groups i.e. glaucoma, cataract and control. Whereas MSs in SP were more consistent between-visit with a statistically significant difference was only found in the cataract group. It indicates the consistency of the SP was not only shown in glaucomatous eyes but in normal subjects as well. In other words, it showed larger between-visit variability of MS with SS as compared to SP. The significant improvement of the MS with SP in cataract group could be due to the much larger proportion of the cataract subjects with reliable results (80%) were perimetric-inexperience prior to their first visit compared to glaucoma (12%) and normal group (33%).

The association of perimetric-experience of subjects with the variability between the visits was found in this study through the comparisons between visits within perimetric-experienced and perimetric-inexperienced subjects. Experienced subjects showed only statistically significant MS improvement in SS whereas inexperienced subjects showed significant MS improvement in both strategies. Experience definitely is a factor which helps to improve sensitivity as it was shown in this study that experienced normal subjects had higher MS either with SS or SP compared to inexperienced normal subjects (pg. 100). The impact of learning effect was minimal with shorter testing duration (Yenice and Temel, 2005). It was shown with SP which apparently had much shorter testing time than SS and it showed no significant difference of MS between visits both with and without perimetric-experience in normal subjects. Whereas SS continued to have significant changes between-visit regardless of experience. On top of that, SP comprises of four phases which the first phase would serve as a training phase to prepare the subject for the definitive examination of DLS estimation (Gonzalez de la Rosa et al., 2013). It could possibly minimize the learning effect of SP.

The lower between-visit variability in SP was also shown through the pointwise analysis of MS between the visits. This pointwise analysis was only targeted among the normal subjects as the recruited glaucoma and cataract subjects had uncontrolled variables in the location and depth of the VF defects which would complicate the data analysis (Bengtsson et al., 1998). The SD which used to represent the variability of the data (Barde and Barde, 2012) was found to increase with eccentricity for the pointwise comparison of MS of SS between visits even though the Mean mean difference of MS between-visit was not statistically significant different among zones according to the eccentricity of the test points. It reiterated the dependency of inter-test variability on eccentricity (Heijl et al., 1989a; Chauhan and Johnson 1999; Wild et al., 1995; Blumenthal et al., 2003). Most of the test points showing higher sensitivity in the second visit were mainly credited to the learning effect except six test points which were mostly located at the peripheral >20° superiorly and temporally. The sensitivity reduction of only six out of the total of 66 test points may not deduce any clinical significance to this study.

SPARK Precision appeared to have smaller SDs of the pointwise variation of MS which indeed is the objective of SPARK was designed for i.e. to reduce fluctuation of the test (Gonzalez de la Rosa et al., 2013). The use of average sensitivity estimation in SPARK was claimed to possibly reduce the fluctuation of the VF test up to 40% in early glaucoma (Gonzalez de la Rosa and Gonzalez-Hernandez, 2011). It was shown smaller changes of MS of SP between-visit compared to SS. Lower inter-test variability could increase the ability to detect statistically significant improvement in MS in the central zones at the second visit compared to in peripheral zones. It was not in agreement with other studies that showed better learning effects at peripheral test points than central (Castro et al., 2008; Heijl et al., 1989; Wood et al., 1987). The lack of between-visit MS improvement in the peripheral test points particularly was partly attributed to the reduction of the sensitivity for a significant number of test points located in the superior peripheral zone. This contradicts the assumption that subjects learned to be more cautious with their upper eyelid in the second visit and eyes would be more opened up during the test (Wood et al., 1987). It could be related to the position of the monitoring camera in Oculus Twinfield which is located lower than the eye fixation level. This could provide false information to the examiner that the upper eyelid does not block the upper edge of the pupil. Therefore, inadequate reminders were given to the subject during the VF test. As the test points in the same zone are correlated in term of threshold sensitivity (Gonzalez de la Rosa and Gonzalez-Hernandez, 2013), almost all the test points in the superior zone displayed the similar form of sensitivity change in the second visit. It was only the superior zone showing reduced average sensitivity in the second visit compared to the other five zones. Further study may be needed to rectify the reduction of sensitivity in the superior zone.

The sensitivity variation between-visit did not correlate with the subjects' age for either strategy which contradicted the findings by Katz and Sommer (1987) but was in agreement with Blumenthal et al. (2003). Standard deviations were used in the two mentioned studies to determine the variability whereas the current study used the mean differences. The fatigue

effect was expected to increase the variability more on the older group (Hudson et al., 1994) but it was not shown in this study with either strategy. It could be the mean age of normal subjects (37.6 years) in this study was much lower than the mean age of the normal subjects (67.2 years) recruited in Hudson et al. (1994). Younger subjects could produce lower inter-subject variability. It might also be due to the fast threshold strategy used in the current study which has shorter testing time and lower variability. All these contributed to the narrower significance level which indeed showed the different statistical outcome.

As the reliability of the VF test could affect the outcome of the result, the reliability indices obtained in the first visit for both strategies were evaluated in this study as well. Even though there was no statistically significant difference in the number of subjects with unreliable results between SS and SP but we found clinically significant more unreliable results with SS compared to SP across all three groups of the subjects: glaucoma (18.1% vs 11.4%), cataract (12.9% vs 6.5%) and control (13.6% vs 2.5%). Most of the unreliable results were mainly due to fixation loss (FL) which was also reported by numerous studies (Rao et al., 2015; Katz & Sommer, 1988; Bickler-Bluth et al., 1989; Nelson-Quigg et al., 1989; Johnson and Nelson-Quigg, 1993; Sanabria et al., 1991; Keltner et al., 2007; Barkana et al.,2006). This could well be attributed to the Heijl-Krakau method used by SS for FL monitoring in which the physiological blind spot is involved whereas the central threshold is used by SP to determine the FL rate. The size and morphology of the optic disc can affect the initial mapping of the blind spot (Phu et al., 2017; Henson et al., 1995) which leads to the higher rate of FL. The difference between strategies was more prominent in the control group which SS had 11 unreliable results (13.6%) compared to only two unreliable results (2.5%) in SP and all the unreliable results in SS were due to FL > 20%. The unreliable results could be possibly avoided with more appropriate instruction (Kutzko et al., 2000) and more careful fixation monitoring especially in the first minute of the examination when the blind spot mapping is initiated. The mapping of the blind spot should be repeated if high FL rate was found at the beginning of the test as recommended by Sanabria et al. (1991). A short training session such as SPARK Training

which is the first phase of the four phases test in SPARK or video training (Sherafat et al., 2003) may be beneficial as well.

It was found in this study that more FN>20% in the glaucoma group compared to the other two groups (four in glaucoma, one in cataract and none in control group). It could be partly due to longer test duration with increased severity of VF defect (Wild et al., 1999; Roggen et al., 2001) and hence induced fatigue effect (Gonzalez de la Rosa and Pareja, 1997; Hudson et al., 1994). The FN rate was not able to be compared between the strategies as it was not available in Oculus perimeters for time-saving measure. Nonetheless, by excluding the subjects with unreliable results due to FN>20%, there was still more unreliable results in SS compared to SP especially for control group which no subject was found to have FN>20% with SS.

False positive (FP) rate was not distinctly different between the two strategies. It could not be compared with any form of statistical method because very few eyes had FP>33% across all three groups for both strategies (One each with SS and SP in glaucoma group; each cataract and control group had one FP>33% with SP). It was the least common index to be encountered compared to the other two reliability indices (FN and FL) and was regarded as the most stable indicator for the reliability (Bengtsson, 2000). The method used in SP to identify FP is determined through the subject's response in catch trials while SS does not use catch trials. SS relies on any subject's response outside of the allowable time (Newkirk et al., 2006; Olsson et al., 1997). High FP rates could affect the values of MD and PSD (Junoy Montolio et al., 2012; Newkirk et al., 2006; Ishiyama et al., 2015) but its effect would minimally influence the results of this study. Moreover, the MD and PSD were not compared in this study. Generally, the proportion of the eyes with unreliable results was higher in the glaucoma group (18.1% in SS) which was comparable to the study reported by Rao et al. (2015) (17.2%) and Rao et al. (2017) (18%) but the FP cut-off rates used were 15% and 20% respectively. However, further evaluation of the data in the study showed the proportion of glaucomatous eyes with unreliable

results remained the same with SS (18.1%) even if the FP cut-off rate of using SS reduced to 15%.

It was also shown that the prior experience in the perimetric test was associated with the reliability of the test when using SS but not with SP in the glaucoma group. The association was not found in the other two groups (cataract and control) even though a higher number of inexperienced subjects was recruited in those two groups (83.9% in cataract and 36.1% in control). There were nine glaucoma patients (20.5%) identified as a novice to the perimetric test prior to the study but they contributed to 45% for the unreliable results in the glaucoma group. This points to the importance of the perimetric experience in achieving better reliability especially for patients with more severe field defect. The learning effect was greater in areas with more severe defects (Kulze et al., 1990; Heijl and Bengtsson, 1996). This resonated well with the significant improvements of MS of SS between visits for the glaucoma group. Even though the learning effect was also reported to present in normal subjects using SITA (Castro et al., 2008; Aydin et al., 2015), as well as in this study (MS of SS statistically significant improved in normal group, pg. 98), deeper field defect and longer testing duration are the two factors causing the perimetric experience in glaucoma group to have a bigger impact in achieving better reliability and improved threshold sensitivity. In addition, visual fatigue was found more in glaucomatous eyes (Heijl, 1977a; Heijl and Drance, 1983; Johnson et al., 1988) which made glaucoma subjects susceptible to unreliable results compared to normal subjects.

Birt et al. (1997) stated that the severity of glaucomatous field defect, duration of the test and age are the factors that affect the reliability of the test. As the same groups of subjects had undergone both SS and SP, this indicates the severity of glaucomatous field defect and age are unlikely to be the reason for the significant association between prior experience in perimetry and reliability of the results found in SS (pg. 93). But the longer test duration in SS, partly due to additional test points in SS (76 of SS vs 66 of SP), maybe one of the reasons for

the association between perimetric experience and reliability of the test in SS. More interval breaks during the VF test would help to improve the results (Hudson et al., 1994).

Testing order was found to have no statistically significant effect in all three groups pertaining to the reliability of the VF test either with SS or SP. Nevertheless, the number of subjects who started with each strategy was well-distributed in all three groups. There was also no effect of the testing order to the reliability of the test according to Barkana et al. (2006) but Barkana et al. (2006) recruited subjects with perimetric experience whereas the current study had a mixture of subjects with and without perimetric experience. However, the first test of subjects with unreliable results was well-divided between SS or SP in this study which should keep the effect of the testing order i.e. the counterbalance between the learning effect and fatigue effect (Brenton et al., 1986) of the subjects to the minimum in this study.

Ultimately according to this study, it is imperative to have a familiarization process before further comparisons between the two strategies are conducted. The learning effect was not only shown in SS but in SP as well. The extent of the effect was different between the two strategies especially in the cataract group in which more than 80% of the subjects recruited were perimetric-inexperienced. Different amounts of variability were found between strategies which provided different impact of the learning effect. Extra visits or training could always minimize the impact and increase the significance of the following comparison study.

3.6 Conclusion

SPARK Precision produced more consistent MS between-visit with only the cataract group showing significant improvement of MS in the second visit, most probably due to a higher rate of perimetric-inexperienced subjects. On the other hand, MS of SS had improved in the second visit for all three groups. Experienced subjects would have higher sensitivity compared to inexperienced subjects regardless of using SS or SP.

Pointwise variation between-visit was significantly smaller with SP compared to SS in normal subjects. SITA Standard showed increased variability with eccentricity but SP showed the absence of MS increment in the superior zone which still needs to be further rectified. Neither of the strategies showed the association of sensitivity variation with subject's age.

The results of this study showed clinically higher rates of unreliable results were found with SS compared to SP, especially in the control group. It was mainly due to high FL which was possibly ascribed to different methods used in FL monitoring and the length of the testing time in both strategies. Different methods were also used in FP rate determination for the strategies but it had minimal effect in the unequal rate of the unreliable results between the two strategies.

The reliability of the test was found to be associated with perimetric-experience of the subject only when using SS in the glaucoma group. It showed the importance of the prior perimetric-experience in achieving a reliable VF test especially in eyes which exhibited field defect. Such association was not found with SP. In contrast, the testing order has no association with the reliability of the test across all three groups (glaucoma, cataract and normal) using either SS or SP.

CHAPTER 4

BETWEEN-STRATEGY COMPARISON BETWEEN SPARK PRECISION AND SITA STANDARD IN NORMAL SUBJECTS

4.1 Introduction

The final threshold value determined in standard automated perimetry (SAP) is based on the concept that 50% of the time the stimulus will be seen by the subject (Schiefer et al., 2005; Weijland et al., 2004). To increase the accuracy of the final value, more attempts should be given and the average value calculated although this would cause the examination time to be inevitably extended. Fatigue would be induced in the extended examination (Gonzalez de la Rosa and Pareja, 1997; Hudson et al., 1994) which could nullify the benefit of multiple attempts. SPARK, a fast threshold strategy incorporated in the Oculus Perimeter (Oculus, Wetzlar, Germany) uses 4 threshold estimates which are determined through 4 different phases to result in an averaged final threshold estimate for each testing point (Gonzalez de la Rosa and Gonzalez-Hernandez, 2013). Its advantage compared to SITA was discussed in the previous chapter showing a lower variability (Chapter 3.4.6, pg. 101).

The shorter testing time in SPARK is achieved by utilising the statistical relationship between the sensitivity values corresponding to the location of test points which are correlated according to six regions (see Figure 1.1 & 1.2, pg. 44) described by Garway-Heath et al. (2000) and Gonzalez de la Rosa et al. (2002). The relationship was established using the results of more than 90,000 VF tests which allows fast probability based estimation of the threshold sensitivity. Description of SPARK was done in more details in Chapter 1.4.2 (pg. 43).

A direct comparison between SITA and SPARK has not been published to date. As SITA is widely regarded as the current “gold standard” in VF test, it is common for a new threshold strategy or VF test to be compared to it. A comparison only between the differential luminance

sensitivities (DLS) produced by SITA Standard (SS) and SPARK Precision (SP) which recalculated mean sensitivities (MSs) were used for the matching points. Usage of the DLS allows the least manipulated data analysis without the influence of each age-matched normative database from the manufacturer which is inaccessible.

4.2 Objectives

The main objective of this chapter was to determine the validity of the SPARK Precision (SP) of Oculus Twinfield 2 in normal sensitivity estimations as compared to SITA Standard (SS) of Humphrey Field Analyzer (HFA). The comparison of these two strategies was achieved through the following specific objectives:

a) To determine the extent of any differences in the group mean of the MS of the matching test points between SS and SP
b) To compare the mean test duration between SS and SP
c) To determine the correlation and agreement between the MS of the matching test points of SS and SP by using the Spearman correlation coefficient and Bland-Altman plot
d) To determine the pointwise bias and limits of agreement (LoA) between the threshold sensitivities of the two strategies at all matching 66 test points using Bland-Altman plots
e) To determine the effect of the eccentricity of the test points and functional zone on the pointwise bias of MS for both strategies

Besides that, a few more objectives were also determined as well; the relationship between mean sensitivity and age, mean sensitivity and spherical equivalent, the bias of mean sensitivities between the strategies and age of the normal subjects.

4.3 Methods

4.3.1 Research Participants

The subjects were recruited from walk-in and existing patients in the optometry clinic, staff and university students. The subjects had given their consent before undergoing a comprehensive optometric examination to determine their eligibility in this study. Only if the subjects fulfilled the following inclusion criteria, they would be asked to proceed to the VF test:

a) Age ranged from 20 to 80 years old
b) BCVA of 6/6 or better
c) Refractive errors below ±6.00DS in sphere and 2.50DC or below in astigmatism
d) No sign of glaucoma or cataract
e) No history of intraocular surgery complication and other ocular diseases that could affect VF
f) No systemic disease that could affect ocular health such as diabetes mellitus and hypertension.

They would be excluded if they presented as pregnant or nursing, taking medications that would affect VF, use of alcohol, nicotine and caffeine less than 2 hours before their eye examination.

The subjects would also be excluded if they failed to perform VF test, poor reliability in their VF results (FP > 33%, FN > 20% and FL > 20%) (Anderson and Patella, 1999) or exhibited VF defect in consecutive visits according to HPA (Hodapp et al., 1993) and Gonzalez de la Rosa (Written communication, 2013). All subjects had completed the VF tests using both SS and SP in their first visit as the familiarization process as shown to be necessary in Chapter 3. The results in the first visit were not included in this study.

4.3.2 Methods and Procedures

The procedure was described in Chapter 3.3.2 (pg. 83). In brief, the eligible subjects underwent VF assessments with the Oculus Twinfield 2 (Oculus, Wetzlar, Germany) using SPARK Precision (SP) and the Humphrey Field Analyzer (HFA) (Carl Zeiss Meditec, Dublin, CA) using SITA Standard (SS) within a day for two visits. The two visits were conducted on different days. The testing order remained identical for the two visits. Only the results produced in the second visit and one eye from each subject were used for the data analysis.

The MSs of the 66 test point of SP and matching points of SS were calculated for each patient and within-visit between-strategies analysis was carried out for the comparison and determination of the relationship between the MS of the two strategies. The left eye results were converted into the right eye for the purpose of the analysis. The association between the MS and age was determined for each strategy and followed the agreement between the MS of both strategies. The test duration of each strategy was recorded and compared between the strategies.

The data analysis in control subjects was continued with pointwise analysis between-strategies across each of the 66 test points by determining the mean difference and 95% LoA of threshold sensitivities. The averages of the mean difference were compared among the six regions according to the functional map described by Gonzalez de la Rosa et al. (2002a) and also among the three regions according to eccentricity (peripheral, mid-peripheral and central).

4.3.3 Statistical Analysis

SPSS version 22 (IBM Corp, Armonk, NY) was used for this between-strategy data analysis. Normality of the data was determined using Shapiro-Wilk test. Wilcoxon Signed Ranks test was used for within-visit between-strategy comparison. Bland-Altman plots (Bland and Altman, 1999) were used to determine the agreement of the MS between both strategies and regression test was continued to determine proportional bias. In Bland-Altman plots, the

differences between the MSs were plotted against the mean and the mean differences between the MSs of the two strategies and 95% limits of agreement (LoA) were determined. All the correlation tests were determined using Spearman's correlation coefficient depending on the normality of the data. One-way ANOVA was used for comparison among the zones and Tukey HSD was used for the post-hoc analysis. A value of $p<0.05$ was considered statistically significant in the comparison or correlation analysis.

4.4 Results

In total, eighty-three healthy control subjects had normal and reliable VF results with both SS and SP in their second visits. There were 49 males and 34 females. Their mean age was 39.6 years (SD 16.1 years; median 36.0 years and range 20 to 71 years). Their median of spherical equivalent was -1.25D (range +1.50D to -6.25D). The testing order was randomly changed among the subjects with 43 subjects started with SS and another 40 started with SP. The means and standard deviations of MS from both strategies calculated from the 66 matching test locations are shown in Table 4.1. The test duration of each strategy was recorded which are shown in Table 4.2.

Table 4.1: The mean sensitivity of SITA Standard and SPARK Precision in normal subjects

	Mean Sensitivity (dB)				
	Mean	SD	Median	Range	
				Min	Max
SS	30.81	1.32	30.94	27.32	33.61
SP	31.74	1.50	31.94	28.06	34.32

Table 4.2: The test duration of SITA Standard and SPARK Precision in normal subjects

	Test duration (minute)				
	Mean	SD	Median	Range	
				Min	Max
SS	6.05	0.67	5.87	4.92	8.18
SP	3.50	0.13	3.50	3.23	3.82

Normality of the data distribution was determined using the Shapiro-Wilk test. It showed the data for age, spherical equivalent, MS of SP and test duration of SS were not normally distributed ($p \leq 0.001$) (See Appendix A4.1).

Direct comparison between the MS of the strategies showed SS had statistically significant lower MS compared to SP (Wilcoxon Signed Rank test: $Z = -6.118$, $p < 0.001$) whereas SP produced statistically significant shorter testing time (Wilcoxon Signed Rank test: $Z = -7.914$, $p < 0.001$) compared to SS.

4.4.1 Relationship between Mean Sensitivities and Age and Spherical Equivalent

The MSs of SS and SP were also shown to be each significantly negatively correlated to subject's age (Spearman correlation coefficient: rho = -0.382, $p < 0.001$ for SS; rho = -0.752, $p < 0.001$ for SP). By using linear regression, a statistically significant (t-test: $t = -4.585$, $p < 0.001$ for SS; $t = -12.470$, $p < 0.001$ for SP) negative slope was found. Mean sensitivity decreases in older subject for both threshold strategies. The slope of the regression line showed -0.037dB/yr for SS and -0.076dB/yr for SP.

Contrarily, no statistical significant correlation was found between the MS from both strategies and the spherical equivalent of normal subjects (Spearman correlation coefficient: $r = 0.166$, $p = 0.133$ for SS; $r = -0.005$, $p = 0.966$ for SP)

4.4.2 Relationship and Agreement between Mean Sensitivities of SITA Standard and SPARK Precision

A statistically significant positive correlation between MS of SS and SP was found for the normal control subjects (Spearman correlation coefficient: rho = 0.713, p < 0.001). Bland-Altman plots were used for the comparison of the MS between the two strategies and it is shown in Figure 4.1. The bias/mean difference (95% LoA) in MS between SS and SP for the normal subjects was 0.92 dB (-1.12, 2.97 dB) (MS of SP higher than MS of SS).

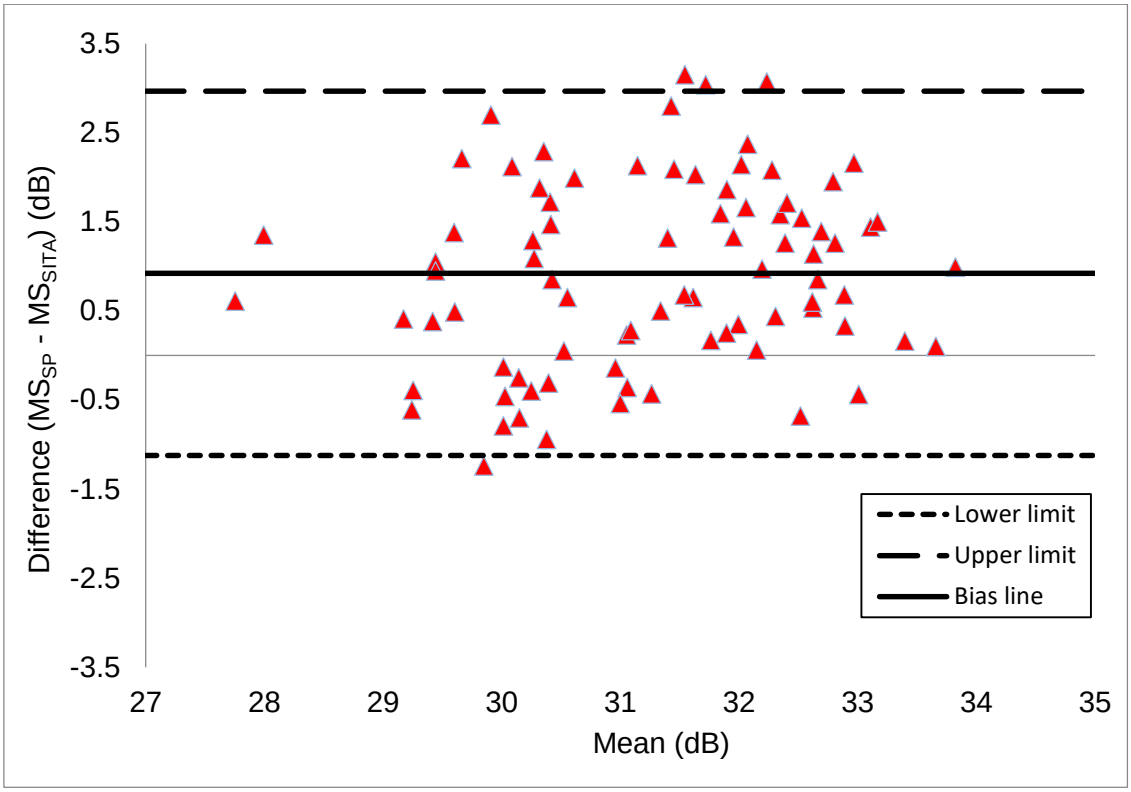


Figure 4.1: Bland-Altman plots for the comparison of the mean sensitivity between SS and SP in control subjects

It was shown that over 95% of the data points (79/83) are located within the LoA and further regression test on the differences showed no statistically significant proportional difference/bias (t =1.713, p = 0.09). Thus, an agreement was shown between the MS of SS

and SP in normal subjects. The mean difference of the MS between the two strategies was also found to have a statistically significant negative correlation to the age of control subjects **(**Spearman correlation coefficient: rho = -0.591, $p < 0.001$). It indicated that the older the subject is, the lesser the difference between MS of SS and SP.

4.4.3 Pointwise Analysis between Strategies

The mean differences (bias) and LoA of pointwise MS between the two strategies for all 66 matching test points were determined. Most of the test points showed higher sensitivity with SP than SS except four temporal points and one inferior point. The mean and SD of the differences of MS between strategies according to the six regions described in the functional map by Gonzalez de la Rosa et al. (2002a) are shown in Table 4.3.

Table 4.3: Means and standard errors of bias in MS between strategies according to six regions described by Gonzalez de la Rosa et al. (2002a)

Region	Bias in MS (dB)	
	Mean	SE
Superior	1.37	0.22
Superior nasal	1.06	0.07
Inferior nasal	1.10	0.20
Inferior	0.43	0.12
Temporal	0.11	0.15
Central	0.92	0.21

A statistical difference was shown across the six regions (One-way ANOVA: $F = 11.192$, $p < 0.001$) which largest mean difference/bias of MS between SS and SP was found in the superior region whereas temporal had the smallest. Post-hoc analysis using Tukey HSD showed

temporal region had the lowest average mean difference compared to the rest of the locations ($p < 0.05$) except inferior ($p = 0.633$).

The pointwise bias of MS between the two strategies was also compared according to the eccentricity of the test points i.e. central < 10°, mid-peripheral 10°-20° and peripheral >20°. The group mean and SD of the pointwise bias of MS in each location according to eccentricity are shown in Table 4.4. There was no statistically significant difference between these three locations (One-way ANOVA: $F = 0.146$, $p = 0.865$).

Table 4.4: Means and standard deviations of bias in MS between strategies according to test point locations

Location	Bias of MS (dB)	
	Mean	SE
Central	0.82	0.08
Mid-peripheral	0.85	0.12
Peripheral	0.91	0.13

If the pointwise bias between-strategy was analyzed according to the range of the bias, it showed that the bias progressively reduced from superonasal to inferotemporal (Figure 4.2). If a vertical midline was used to equally divide the nasal and temporal test locations, nasal test points significantly had higher average bias than the temporal test points (Unpaired t-test: $t = 5.664$, $df = 64$, $p < 0.001$) (Table 4.5).

Figure 4.2: Bias (upper bolded number) and the 95% limits of agreement (lower unbolded number) of sensitivity values for 66 matching test points between SS and SP

* Bias = MS_{SP} - MS_{SS}
*The colours represent the range of the bias of sensitivity values between-strategy

Table 4.5: Means and standard deviations of bias in MS between strategies according to areas divided by vertical midline

Location	Bias of MS (dB)	
	Mean	SE
Nasal	1.21	0.07
Temporal	0.51	0.10

4.5 Discussion

In this study, a group of normal subjects with a wide age range was recruited to determine the performance of SP against an established threshold strategy, SS. The participants first visit VF results were disregarded in this study which was important (for further details see Chapter 3.4.4.3, pg. 96) to minimize the learning effect.

The MS of SP was significantly higher than that of SS (0.92dB) while a reasonable agreement between the two strategies in normal subjects was found as supported by the Bland-Altman test. According to the evaluation criteria suggested by Luithardt et al. (2015), a bias of less than 1dB indicates a good agreement but LoA of more than 4dB has demoted the agreement to “acceptable” [LoA in this study was reported as (-1.12, 2.97 dB), pg. 121]. All 66 matching VF test points were used for the comparison of MS which is the least manipulated index without the influence of age-matched normative database from each strategy. The higher average threshold estimates produced by SP could partly be due to reduced fatigue effect with shorter test duration. Test duration of SS in this study was around six minutes for all 76 test points whereas SP required three and a half minutes as an average for 66 test points. Bengtsson et al. (1998) also reported similar test duration for SS whereas SP was targeted to complete the test to about half of the time (Gonzalez de la Rosa and Gonzalez-Hernandez, 2013). By referring to Gonzalez de la Rosa and Pareja (1997), a reduction rate of 0.08 to 0.1 dB in MS per minute was associated with fatigue effect. The additional testing time of 2.5 minutes with SS was possibly accountable for up to 0.25 dB reduction of MS which just partly explained the disparity in threshold estimates between SS and SP. SITA Standard produced 50% shorter time compared to Full Threshold (FT) (Bengtsson et al., 1998) observing 0.8 to 1.9 dB higher MS (Wild et al., 1999; Artes et al., 2002; Bengtsson et al., 1998). Many studies had shown longer testing times produced lower threshold sensitivities in normal subjects (Searle et al., 1991; Heijl, 1977a; Hudson et al., 1994; Johnson et al., 1988; Langerhorst et al., 1987; Suzumura, 1988; Wildberger and Robert, 1988) which was also shown when using fast threshold strategies (Bengtsson et al., 1998; Wild et al., 1999; Glass et al., 1995; Flanagan et

al., 1993; Schaumberger et al., 1995; O'Brien et al., 1994; Nordmann et al., 1994; Morales et al., 2000). Therefore, a fast threshold strategy that produced shorter testing time such as SPARK strategy may produce threshold estimates that are closer to the "real" sensitivity value (Gonzalez de la Rosa and Gonzalez-Hernandez, 2011) which is relatively less affected by fatigue effect.

The differences in MS between the strategies could also be related to the methodology applied in each algorithm. SITA utilizes two likelihood functions developed from prior knowledge of normal and glaucomatous models which continue to be adjusted according to the patients' response during the test (Bengtsson et al., 1997). It uses step size of 4-2 dB to determine the sensitivity value that achieving 50% probability of seeing which is terminated when predetermined level of accuracy specified by ERF is obtained (Olsson and Rootzen, 1994) and recalculation of the final sensitivity using postprocessing algorithm is conducted for each stimulus point at the end of examination (Bengtsson et al., 1997; Bengtsson et al., 1998). On the other hand, the initial stimulus for the first phase in SPARK is referred to a sample of examinations comprised of 5% of cases with neurological, retinal and mixed diseases besides glaucomatous and normal eye results (Gonzalez de la Rosa and Gonzalez-Hernandez, 2013). The information from the non-glaucomatous cases was claimed to be helpful in determining the MD but not accurate in detecting the irregularity of the VF (Gonzalez de la Rosa and Gonzalez-Hernandez, 2013). SPARK uses the first phase stimulus to determine the threshold estimates for the other three phases. Interpolation and multiple regression equations are used in determining the threshold estimates. The new sensitivity value is estimated using the previous sensitivity value plus or minus the standard error (SE) depending on the response of the subject. The SE is used as a step to refine the deviation of each test point for the following phase. It is repeated for the deviation estimation for all the points within the same previously defined sectors (Gonzalez de la Rosa et al., 2002). A median or mean of the threshold estimates from the 4 phases is calculated for each stimulus point and excluding the most extreme threshold estimate could minimize the errors possibly made during the process. The

question of which strategy produces the threshold estimates which is closer to the “real’ values needs to be further determined regardless of the fatigue effect.

The present study also showed that the MS of SP and SS is found to have significant associations with age ($p < 0.001$ for both strategies). It decreased with age which in agreement with previous reports (Brenton and Phelps, 1986; Haas et al., 1986; Jaffe et al., 1986; Heijl et al., 1987; Iwase et al., 1988; Spry and Johnson, 2001). The reduction rates were approximately -0.4dB/decade and -0.8dB/decade for SS and SP respectively which are close to the rates reported by Brenton and Phelps (1986) and Weijland et al. (2004) (-0.6dB/decade) and also Wild et al. (1999) and Heijl et al. (1987) with rate about -0.7dB per decade.

Ageing was also found to be associated with the reduction of the bias between the strategies. As mentioned earlier that the rate of sensitivity reduction relative to age was higher with SP (-0.08 dB/yr) than with SS (-0.04 dB/yr). The age-matched sensitivity values published by the manufacturer (Weijland, 2004) were used in the SP algorithm which had the predetermined rate of the sensitivity deterioration against age. The threshold sensitivity estimated from SP was higher in younger normal subjects. The higher deterioration rate would progressively decrease the bias between the strategies towards older age groups. Wild et al. (1999) showed the mean difference of threshold estimates between SS and FT which was independent of age. The consistency could be due to the fact that both strategies are from the same manufacturer (Carl Zeiss Meditec, Dublin, CA, USA) which possibly refer to the same set of normative database.

Oculus Twinfield perimeter was reported to produce 1.5 dB higher DLS compared with HFA using FT strategy in normal subjects (Lorch et al., 2001) but specifications of the Twinfield was not provided or whether any data conversion was performed. Capris et al. (2008) showed that Twinfield produced lower MS using FT strategy due to the maximum stimulus luminance used in Twinfield was 1000 asb (318 cd/m^2) compared to 10,000 asb in HFA. Humphrey Field

Analyser and Oculus Twinfield have some similarities in the characteristics of instrument setting such as background luminance, stimulus size, stimulus duration and working distance. The stimulus delivery method is different between the perimeters. Projected stimuli are used in HFA whereas Oculus Twinfield uses a semi-transparent bowl which allows back-projected stimuli to be observed. Due to the difference in maximum stimulus luminance, the minimum threshold value of 0 dB in Oculus Twinfield was basically equivalent to 10 dB in HFA (Capris et al., 2008). The average difference between the MS of FT strategy of Oculus Twinfield 2 and HFA was 8.90 dB (Capris et al., 2008) which was less than the theoretical calculation. Wabbels et al. (2001) also found that Oculus FT strategy produced higher MS of 0.5 dB than SS in normal subjects after using corrected values for different maximum luminance. Considered the difference between the mean difference of 0.92 dB found in this study and the gains of up to 0.25 dB by saving 2.5 minutes in testing time, perhaps there was another instrumentational factor from Oculus Twinfield which could still contribute to the higher threshold estimation by SP. Different staircase algorithms and method in threshold estimation were claimed to be responsible for the difference in MS between Oculus and HFA. Nevertheless, the current version of Oculus Twinfield 2 is also only able to produce up to 1000 asb for the maximum stimulus intensity. Simulated maximum stimulus luminance up to 10,000 asb is artificially achieved when using SPARK strategy by modifying the reference luminance level used to obtain the decibel scale in HFA. Hence, the corrected threshold value was not required in this study.

It was worth to note that the background light source used in Oculus Twinfield was halogen light but HFA used fluorescent light. Both exhibited different characteristics of the spectral distribution of white light which the background illumination of the Oculus Twinfield apparently appeared to be more yellowish or warmer than HFA's. It showed lower colour temperature in Oculus Twinfield as compared to HFA) (grossly 2395K vs 4721K without strict control of total darkness of the room during the measurement using Chroma meter CS-150, Konica Minolta. Inc, Osaka, Japan). Bright lights give a more pleasant feeling whereas warm lighting lets

patients feel more relax (Park et al., 2013). Nevertheless, there was no difference in subjective assessment of fatigue (Janosik and Marczak, 2016) between in lighting with higher colour temperature (cool) and lower colour temperature (warm). Lou et al. (2011) also showed there was no difference of visual fatigue under three different colour temperatures from 2700K, 6500K and 10000K which 40 minutes of reading task was regarded to be too short to induce significant visual fatigue. Each perimetric tests either with SS or SP were completed within a much shorter time. Moreover, it was suggested that only lighting of more than 5000K would provide better visual performance especially for elderly (Yamagishi et al., 2008) whereas the background illumination of HFA was lower than that. Hence, the effect of the comparatively warmer background illumination in Oculus Twinfield was believed to be clinically negligible in this study.

SPARK Precision uses almost identical stimulus locations compared to SS 30-2 except the uppermost and bottommost row of test points which enabled more straightforward pointwise comparison between strategies. Pointwise comparison of MS between strategies showed most of the test points had higher sensitivity values with SP. Only five stimulus locations had higher value with SS than SP and majority of them were located at temporal zone. The temporal region was also the region with the lowest mean bias among the six regions described by Gonzalez de la Rosa et al. (2002a). Its reason remains unknown. On the other hand, the superior region had the highest bias. It was also shown in the previous chapter that the superior region had a different range of variability compared to the other regions (see Chapter 3.4.6, p. 99). It could indicate that the test points located in the superior region are more vulnerable when using SP. As most of the test points could only be tested once and the sensitivity values within the same region are highly correlated, any response error at any test point within the same region could affect the final estimation of the sensitivity value. It was demonstrated in this study that the differences were found among the functional zones rather than being a result of eccentricity. But whether the superior region was affected by eyelid or surveillance camera position still needs to be further investigated.

There was no significant effect of eccentricity to the bias between SS and SP even though SS exhibited greater inconsistent sensitivity at more peripheral regions (Wild et al., 1999). The omission of the uppermost and bottommost test points in the comparison of MS between strategies perhaps had downplayed the influence of the eccentricity. By categorizing the ranges of the pointwise sensitivity biases between the strategies, a trend of bias reduction was observed from superonasal to inferotemporal which also distinguished a higher mean bias in the nasal field compared to temporal. The reason for the changes is yet to be investigated.

The MSs of both SS and SP were independent of mean spherical equivalent (MSE) in this study. This could be due to the fact that subjects with high myopia > -6.00DS and high astigmatism >-2.50DC were excluded. Besides, an optical correction was provided during VF testing for all subjects according to the manufacturers' recommendation. Visual field testing in the present study was conducted for the central 30-degree field which required optical defocus to be corrected especially when Goldmann size III stimulus or smaller was used for both strategies (Atchison, 1987; Anderson and Patella, 1999).

A stratified equal number of subjects in each decade of age could be more ideal for the in-depth analysis of the differences between the strategies. The current study was skewed towards younger subjects with a median age of 36.0 years old (40% was younger than 30 years old). The strict inclusion criterion of BCVA was used for the recruitment of subjects which increased the difficulty of recruiting elderly subjects with healthy and normal vision. All normal subjects recruited in this study were able to achieve minimal BCVA of 6/6. Elderly subjects who have reduced vision were either rejected or recruited for the study described in another chapter if cataract was found (see Chapter 6).

Thirty out of 83 subjects (36.1%) had no experience of VF testing and even though they all had gone through the first visit for the familiarization process and improved results were still

shown in the second visit. As the learning effect could be displayed up to five times of VF test (Wood et al., 1987), it is possible that the improvement of the second VF test was partly due to the learning effect.

4.6 Conclusion

The between-strategy comparison was the extension study from Chapter 3 which the results of the second visit was analyzed and compared within the normal subjects. An acceptable agreement (bias < 1dB, 4 dB < LoA < 5dB) was found between the threshold estimates from both strategies with the mean threshold estimates from SP was higher by 0.92dB in normal subjects. It may be partly due to the reduced fatigue effect in shorter test duration with SP which saved more than 40% of the testing time compared to SS. The differences between methods used in the algorithms and other factors related to the setting of the instrument may also have contributed to the higher threshold estimates in SP. The bias between the strategies was also found to decrease with age.

Threshold estimates from SS and SP showed significant association with age but not with the spherical equivalent of the subjects. SPARK Precision showed a higher rate of threshold sensitivity deterioration against age compared to SS.

Pointwise comparison between the strategies showed the bias was different among the six functional regions described by Gonzalez de la Rosa et al. (2002a) but no difference was found with eccentricity. The ssuperior region had the highest mean bias while temporal was the lowest. A trend of sensitivity bias reduction from superonasal to inferotemporal was also observed. The reason behind the difference among regions or the trend of changes needs to be further investigated but it may indicate that the sensitivity value in each stimulus point is vulnerable to any error made in any test point within the same zone when using SP.

CHAPTER 5

COMPARISON BETWEEN SPARK PRECISION AND SITA STANDARD IN GLAUCOMA PATIENTS

5.1 Introduction

Glaucoma is one of the leading causes of blindness worldwide (Quigley and Broman, 2006; Bourne et al., 2013; Congdon et al., 2004). It is a chronic progressive optic neuropathy that displays characteristic structural changes of the ONH which are often associated with VF loss (Otarola et al., 2016; Weinreb and Khaw, 2004; Fechtner and Weinreb, 1994). To this date, static threshold automated perimetry is still the indispensable tools in clinical practices for the clinical visual function assessment in glaucoma (Phu et al., 2017; Jampel et al., 2011; Nouri-Mahdavi et al., 2011) despite the introduction of advanced imaging techniques such as OCT and HRT particularly for the diagnosis in the early stage of the disease (Kanamori et al., 2006; Lisboa et al., 2012; Fanihagh et al., 2015; Mardin et al., 1999).

New threshold algorithms have been introduced to improve the evaluation of achromatic perimetry in glaucoma patients (Bengtsson and Heijl, 1998; Bengtsson and Heijl, 1998a; Gonzalez de la Rosa et al., 1997; Schiefer et al., 2009). The new algorithms provide shorter testing time which helps to reduce the fatigue effects and intra- and inter-test variability thus improves the diagnostic sensitivity to detection VF changes. SPARK, one such new algorithm (Gonzalez de la Rosa and Gonzalez-Hernandez, 2013) was introduced with the aim to produce a more stable result as shown in the previous study (see Chapter 3, pg. 104) with reduced variability between visits compared to SS. Although it produced slightly higher threshold sensitivities than SS, it displayed acceptable agreement between the sensitivity values estimated by SS and SP in normal subjects (see Chapter 4, pg. 121). Estimation of the sensitivity values using SPARK is designed according to the relationship among stimulus points within the morphologically and functionally defined regions (Gonzalez de la Rosa and

Gonzalez-Hernandez, 2013). Pointwise sensitivity comparison between SP and SS showed the differences varied among the regions but these differences with average less than 1 dB are most likely clinical insignificant for a normal subject (see Chapter 4, pg.121). Therefore, with the consistency and shorter testing time shown in normal subjects, the use of SPARK should be further evaluated in eyes with VF defects particularly glaucomatous eyes that have displayed large variability and extended testing time in VF testing (Chauhan et al., 1993; Flammer et al., 1984a; Gardiner et al., 2012; Aoki et al., 2007; Wild et al., 1999). If SPARK remained to produce shorter testing time and uphold consistency results in the glaucomatous eyes, it could possibly have a better capability in glaucoma diagnosis and monitoring.

5.2 Objectives

This study was a continuation of work from the previous chapters (Chapter 3 and 4) to further evaluate the performance of SPARK Precision from Oculus Twinfield 2 against universally recognised gold standard threshold strategy, SITA Standard of HFA in assessing eyes with VF defect. Glaucoma is widely known to be associated with localised VF loss (Sihota et al., 2007; Steele and Spry, 2009) which has been commonly diagnosed and monitored through VF testing. This study was aimed to compare the efficiency of SP with SS to assess the VF of glaucoma patients and a group of age-matched healthy normal subjects as control.

The aim of this study was achieved through the following specific objectives:

a) To determine and compare the mean sensitivity of the matching 66 test points between the threshold strategies, SS and SP in glaucoma patients and also in age-matched normal subjects
b) To compare the global indices and test duration between the two strategies within each group
c) To determine the agreement of the group MS and global indices between the strategies using Bland-Altman plots within all the subjects, glaucoma and age-matched normal group respectively

d) To determine the correlation of the global indices between the strategies for all the subjects
e) To conduct pointwise comparison of threshold sensitivities between the two strategies using Bland-Altman analysis for the matching 66 test points
f) To determine the extent of any differences of the severities of VF defects between both strategies in glaucoma patients according to the Advanced Glaucoma Intervention Study (AGIS) severity scale (AGIS, 1994)
g) To determine and compare the size and depth of the glaucomatous field defect between both strategies using criteria recommended by HPA (Hodapp et al., 1993)
h) To determine the sensitivity and specificity of SP in assessing VF abnormalities against the results of SS and also to compare the sensitivity and specificity between both strategies according to each recommended criterion.

5.3 Methods

5.3.1 Research Participants

Similarly, as described in Chapter 3.3 (pg. 82), all glaucoma patients were either open or closed angle glaucoma who were medicated and confirmed diagnosis by their ophthalmologist before enrolled in this study. All the glaucoma patients had typical glaucomatous cupping of the optic disc and/or VF loss on 30-2 or 24-2 HFA SS VF testing while patients with a family history of glaucoma and/or suspicious discs but no significant structural changes and normal IOP and VF were recruited as glaucoma suspect in this study. Normal control subjects had healthy eyes with normal fundus appearance and no family history of glaucoma. The detailed criteria were similarly described in Chapter 3 and 4.

Some important inclusion criteria for the subjects recruited in this study are as follows:

a) Aged from 20 to 80 years
b) Best corrected visual acuity (BCVA) 6/12 or better for glaucoma patients and 6/6 or better for normal subjects

c) Refractive errors below 6DS in sphere, 2.5DC or less in astigmatism.
d) No history of intraocular surgery complications and no other ocular disease besides glaucoma and systemic illness that could affect VFs
e) No uncontrolled diabetes mellitus or untreated hypertension
f) Not pregnant or nursing
g) Not taking any drugs or alcohol that potentially affects reaction time or VF

The recruited subjects must have acceptable reliable indices from both of their VF tests using SS and SP (FP ≤ 33%, FN ≤ 20% and FL ≤ 20%) (Anderson and Patella, 1999). Subjects recruited as control initially but exhibited significant VF defect in the second VF test using SS were excluded from this study.

5.3.2 Methods and Procedures

The details of the procedure were similar to those as described in Chapter 3. All normal subjects had undergone a standard routine eye examination to confirm their eligibility for the study.

All glaucoma patients and normal subjects had attended two visits of the VF tests. During each visit, VF tests were performed using SP and SS. A short break of at least 10 minutes was given between the VF tests. The same order of VF tests was applied in the second visit for the same subjects. The same procedure was carried out as mentioned in Chapter 4, the first visit was used as a familiarization process and only the second visit results were used for the comparative analysis between the strategies.

Only the result from one eye of each patient was used for the data analysis. If both eyes were diagnosed with glaucoma, the eye exhibited less glaucomatous field defect was chosen whereas the normal eye was randomly chosen if both eyes of the same patient were also eligible for the study.

The mean sensitivity (MS) values were calculated according to the matching 66 test points between SS and SP. The 30-2 test pattern was used for SS but the uppermost and bottommost rows of test points and also the two test points located at the blind spot were excluded for the calculation of MS. The left eye results were converted into the right eye for the purpose of the analysis. The calculated mean sensitivity and the existing global indices included in the printout i.e. mean deviation and pattern standard deviation from each strategy test were recorded and analysed as well as the average testing time of each strategy test. The number of abnormal pattern deviation points (NAPDP) with at least $P < 5\%$ was also identified for each strategy. The abnormal PD points for uppermost and bottommost test points in SS were not included.

The severity of VF defects among glaucoma patients from both strategies was also determined and compared using the Advanced Glaucoma Intervention Study (AGIS) severity scale (The AGIS investigators, 1994). This scale is based on the number and depth of neighbouring depressed test locations on the total deviation plot of single field analysis in the nasal area, upper and lower hemifield. A test point is considered as a depressed test location when a minimum deviation from normal is reached. The minimum deviation required in each test point is shown in Figure 5.1.

Figure 5.1: AGIS visual field test scoring map

Source: The Advanced Glaucoma Intervention Study (AGIS) investigators, 1994.

The AGIS score ranges from 0 to 20. Its scoring method is shown below:

i) Nasal defect is a cluster of three or more adjacent depressed test locations in nasal field area which may cross the horizontal midline.

ii) Nasal step is formed by one or more depressed test locations in the nasal field, either below or above the horizontal midline, in the absence of depression of any of the three test locations on the opposite side of the horizontal midline.

iii) Hemifield defect is a cluster of three or more depressed sites in a hemifield. More than one cluster of depressed sites may occur in a hemifield.

Its scores were awarded according to Table 5.1.

Table 5.1: The AGIS scoring method

Type of defect found	Score awarded
Nasal defect or nasal step	+1
Four or more of the six nasal test locations are depressed 12dB or more	+1
In each hemifield with one or more clusters of three or more adjacent depressed test locations (hemifield defects), if the number of depressed test sites in the cluster is	
- 3 to 5	+1
- 6 to 12	+2
- 13 to 20	+3
- >20	+4
If half or more the adjacent defective locations in a hemifield are depressed	
- 28dB or more	+5
- 24dB or more	+4
- 20dB or more	+3
- 16dB or more	+2
- 12dB or more	+1
If a hemifield lacks a cluster of three adjacent depressed test sites but contains at least two adjacent depressed of which one is depressed 12dB or more	+1
Maximum possible score	20

Source: AGIS, 1994

The size and depth of the glaucomatous field defect detected by the VF test using SS and SP were determined and compared. The size of the field defect is the number of the points in pattern deviation map which fulfilled one of the criteria of abnormal VF recommended by HPA (Hodapp et al., 1993) i.e. a cluster of three or more non-edge points in either hemifield on

pattern deviation probability map of SS 30-2 with sensitivity found in <5% of the normal population (p < 5%) with at least one of the defective points having p < 1%. The two points which are the far nasal points that one located above and the other one below the horizontal meridian are included as the non-edge points. The matching test points in the SP were used for the calculation of the glaucomatous field defect size. The depth of the glaucomatous field defect is the sum of the sensitivity values from the points was used to identify the size of the VF defect in pattern deviation map (Budenz et al., 2002a; Aoki et al., 2007). The average depth of each abnormal point is the ratio between the size and total depth of the glaucomatous scotomas was determined and compared between the strategies.

The sensitivity and specificity indices of SP in assessing VF abnormality against SS results and also against the diagnosis of glaucoma made by local ophthalmologists were determined. The sensitivity in detecting different severity level of glaucoma according to the AGIS score was also determined for each strategy. Sensitivity is defined as the ability of the threshold strategy to detect glaucoma in a confirmed glaucoma patient whereas specificity is the ability of the threshold strategy to produce a normal result in a confirmed normal subject (Elliot, 2007). The criteria used to determine abnormal VF in SS was according to the criteria recommended by HPA (Hodapp et al., 1993).

A subject was considered having abnormal VF if either one of the following criteria was found in his/her SS results:

a) A cluster of three or more non-edge points in either hemifield on pattern deviation probability map with sensitivity found in <5% of the normal population (p<5%) with at least one of the defective points having p<1%,
b) Pattern standard deviation (PSD) had a value found in <5% of the normal population (p<5%),
c) Glaucoma hemifield test (GHT) showed “outside normal limits”.

Whereas abnormal VF with SP was determined based on the criteria recommended by Gonzalez de la Rosa et al. (2013) which at least 95% specificity was achieved (Written communication, Gonzalez de la Rosa, 2015):

a) Mean deviation (MD) with a value lower than -2.3 dB,

b) Pattern standard deviation (PSD) with a value more than 1.8 dB,

c) More than 5 defect points/scotomas with total deviation> 5dB.

5.3.3 Statistical Analysis

All the data were analysed using SPSS version 22 (IBM Corp, Armonk, NY) and normality of the data was determined by using the Shapiro-Wilk test. Unpaired t-test or Mann Whitney test was used for comparison between-group within-strategy whereas paired t-test or Wilcoxon Signed Ranks test was used for the comparison between-strategy within-group. Bland-Altman plots were used to determine the agreement of the global indices between both strategies and regression test was used to determined proportional bias. The analysis of pointwise sensitivity between-strategy was also conducted with the mean differences and LoA of all 66 test points between the strategies were determined. All the correlation tests were determined using Spearman's correlation coefficient for non-normally distributed continuous variables.

5.4 Results

A total of 39 subjects were recruited as glaucoma patients, comprising of 34 POAG or NTG and 5 PACG. An additional of five subjects were recruited: glaucoma suspects who were included only in the Bland-Altman analyses of this study. All subjects had undergone two visits of VF testing [duration between visits: median (Range) = 7.0 (1 – 35) days] and only the reliable second visit results were used for the analysis in this chapter. There were 30 out of 34 subjects with POAG/NTG and three out of five PACG exhibited VF defect using SS according to HPA criteria (Hodapp et al., 1993). Nineteen patients were females and 20 were males. The same order of the test in both visits was used for each subject i.e. 21 tested first with SS and another

18 started with SP. The distribution of the severity level of glaucoma for the recruited 39 glaucoma patients that were classified according to AGIS score (The AGIS investigators, 1994) is shown in Table 5.2. More than 70% of the recruited glaucoma subjects were categorized in “None” and “Mild” stages of glaucoma. The rest of them were almost equally distributed between “Moderate” and “Severe” stages of glaucoma. None of the recruited glaucoma patients was an “End-stage” patient.

Table 5.2: Severity of glaucoma according to AGIS score for glaucoma subjects

AGIS score	Category	No of patient	Percentage
0	None	16	41%
1 - 5	Mild	12	31%
6 - 11	Moderate	6	15%
12 - 17	Severe	5	13%
18 - 20	End-stage	0	0%
	Total	39	100%

Forty-five age-matched healthy normal controls (25 females) showed no VF defect with SS. They had completed the same test routine as glaucoma patients (two visits of VF test with a similar duration between the visits) [median (range) = 7.0 (2 – 35) days]. The controls were matched not only for age but also for spherical equivalent with the recruited glaucoma group. There were 23 of them started with SS and 22 started with SP for the test order during both of their visits. Comparison of the age and spherical equivalent between both groups is shown in Table 5.3.

Table 5.3: Comparison of age and spherical equivalent between glaucoma and normal group

	Glaucoma (n = 39)				Normal (n = 45)			
	Mean	SD	Median	Range	Mean	SD	Median	Range
Age (yr)	54.7	12.6	55.0	(24–78)	51.9	11.7	51.0	(26–71)
Spherical equivalent (D)	-1.66	2.46	-0.75	(-6.63–3.25)	-1.46	1.79	-0.75	(-5.00–1.50)

Shapiro-Wilk test was used for the data normality test which showed data for age was normally distributed but not for SE (See Appendix A5.1).

Neither age (Unpaired t test: $t = 1.077$, $df = 82$, $p = 0.285$) nor SE (Mann-Whitney: $Z = -0.233$, $p = 0.815$) was found to have statistically significant difference between glaucoma group and normal subjects.

5.4.1 Comparison between Glaucoma and Age-matched Control Group

The global VF indices (MS, MD and PSD), numbers of abnormal PD points and test duration of each strategy (SS and SP) were compared between glaucoma and age-matched healthy control group (Table 5.4). Shapiro-Wilk test was used to determine the normality of the data distribution of all the descriptive data (See Appendix A5.2). Only the test duration for the SP was normally distributed hence unpaired t-test was used for comparison between the two groups. All other comparisons were conducted with the Mann-Whitney test.

Table 5.4: Comparison of VF data between glaucoma and age-matched control group

		Glaucoma				Normal				p
		Mean	SD	Median	Range	Mean	SD	Median	Range	
MS (dB)	SS	25.76	5.17	27.36	(11.24–31.00)	30.33	1.36	30.41	(27.32–33.32)	<0.001
	SP	26.13	5.56	28.42	(11.55–32.26)	30.79	1.29	30.85	(28.06–33.92)	<0.001
MD (dB)	SS	-4.00	4.72	-2.31	(-16.65–0.83)	0.47	1.15	0.75	(-2.21–2.38)	<0.001
	SP	-1.16	3.97	0.32	(-11.14–3.49)	2.58	1.20	2.82	(-2.22–3.96)	<0.001
PSD (dB)	SS	5.86	4.33	4.85	(1.47–14.79)	1.65	0.29	1.64	(1.17–2.54)	<0.001
	SP	3.02	2.43	1.76	(0.94–8.05)	1.16	0.22	1.16	(0.77–1.72)	<0.001
NAPDP	SS	16.2	12.6	14.0	(1 – 47)	3.07	2.78	2.00	(0 – 11)	<0.001
	SP	12.7	13.3	8.0	(0 – 35)	1.18	1.76	0.00	(0 – 6)	<0.001
Test time (min)	SS	7.66	1.55	7.27	(5.47–12.15)	6.32	0.72	6.13	(5.05–8.18)	<0.001
	SP	3.57	0.15	3.55	(3.33–3.87)	3.48	0.14	3.45	(3.23–3.82)	*0.003

*Unpaired t-test
The rest use Mann-Whitney test
NAPDP – Number of abnormal pattern deviation points

There were statistically significant differences between glaucoma and control group with regards to MS, MD, PSD, NAPDP and testing time. Testing duration for SP of glaucoma patients was statistically significant longer than for controls ($t = 3.043$, $df = 82$, $p = 0.003$) even though the nominal difference in the average duration time was less than 6s.

5.4.2 Comparison Between-strategy Within-visit

The global indices, NAPDP and test duration in both groups were compared between the SS and SP (Table 5.5). Paired t-test or Wilcoxon Signed Rank test was used for the statistical comparison. There were statistically significant differences between SS and SP in MD, PSD,

NAPDP and test time for both glaucoma and control subjects (see Table 5.5, pg. 145). The mean MS for normal subjects showed statistically significant higher value for SP compared to SS but the difference was clinically negligible with a value less than 0.5 dB. Whereas MS of glaucoma subjects appeared to have larger differences between the strategies but the difference itself was not statistically significant ($p > 0.05$). Mean deviation (MD) obtained by SP showed much higher values than of SS in both glaucoma and control subjects. Pattern standard deviation (PSD) was distinctly higher with SS than with SP in glaucoma subjects but the difference was much lesser in normal subjects. SITA Standard detected statistically more abnormal pattern deviation points in both groups as compared to SP but the difference was negligible in normal subjects. The testing time of SS was recorded as the testing time for all 76 test points which need more than double of the time for SP to complete the test for 66 stimulus locations in glaucoma patients. It had also saved about 40% of testing time for normal subjects when using SP compared to SS.

Table 5.5: Comparison between SITA Standard and SPARK Precision

		SITA Standard		SPARK Precision		Paired t-test/Wilcoxon Signed Rank test		
		Mean^/ Median	SD^/ (Range)	Mean^/ Median	SD^/ (Range)	t^/Z	df	p
MS (dB)	G	27.36	(11.24-31.00)	28.42	(11.55-32.26)	-0.754		0.451
	N	30.33^	1.36^	30.79^	1.29^	-3.189^	44	0.003*
MD (dB)	G	-2.31	(-16.65-0.83)	0.32	(-11.14-3.49)	-4.996		<0.001
	N	0.75	(-2.21-2.38)	2.82	(-2.22-3.96)	-5.757		<0.001
PSD (dB)	G	4.85	(1.47-14.79)	1.76	(0.94-8.05)	-5.443		<0.001
	N	1.65^	0.29^	1.16^	0.22^	11.364^	44	<0.001*
NAPDP	G	14.0	(1 – 47)	8.0	(0 – 35)	-2.111		0.035
	N	2.00	(0 – 11)	0.00	(0 – 6)	-4.151		<0.001
Test time (min)	G	7.27	(5.47-12.15)	3.55	(3.33-3.87)	-5.443		<0.001
	N	6.13	(5.05-8.18)	3.45	(3.23-3.82)	-5.842		<0.001

G - Glaucoma; N - Normal
*Paired t-test
The rest use non-parametric Wilcoxon Signed Rank test

5.4.3 Agreement and Correlation Between-strategy

5.4.3.1 Mean sensitivity

The means MS from the matching 66 test points for both glaucoma and control groups were compared between the strategies. There was a statistically significant correlation between the MS of SS and SP (Spearman correlation coefficient: rho = 0.804, $p < 0.001$). Bland-Altman plots were also used to determine the agreement of threshold estimates between the strategies. Figure 5.2 shows the Bland-Altman plot of MS between SS and SP for all the glaucoma patients, glaucoma suspects and control group.

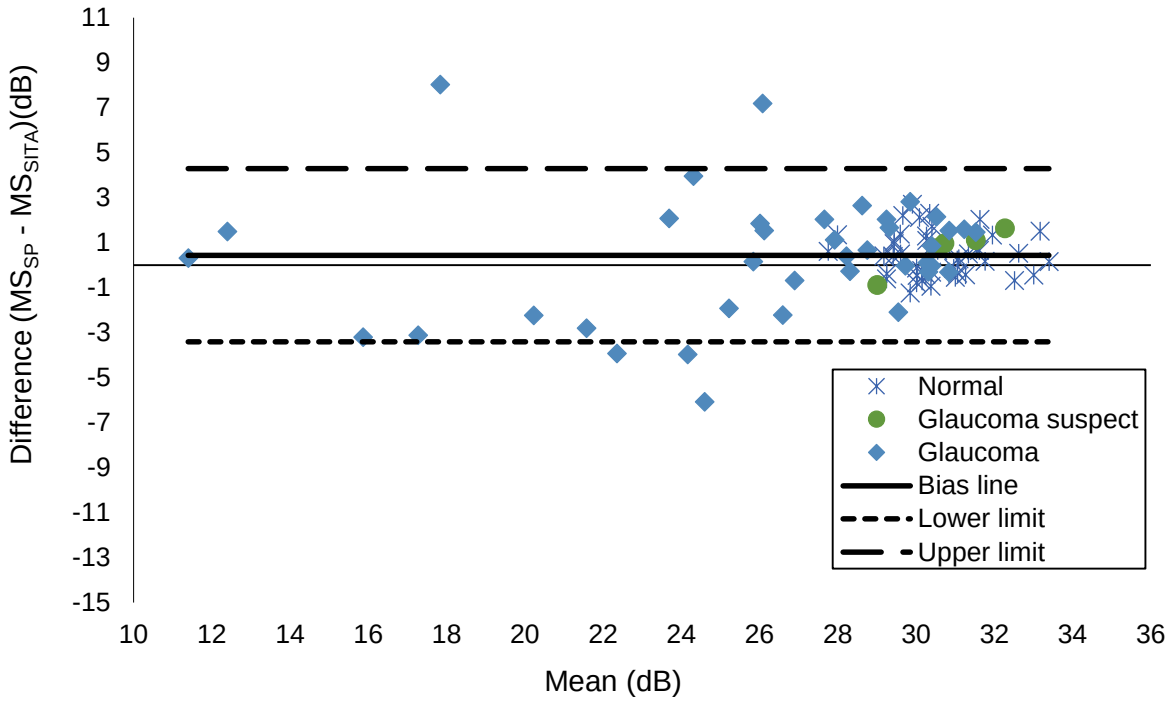


Figure 5.2: Bland-Altman plot of mean sensitivity between SS and SP in glaucoma patients, glaucoma suspects and age-matched normal subjects

The bias/mean difference (95% LoA) in MS between SS and SP for all subjects including the glaucoma suspect was 0.44 dB (-3.41, 4.29 dB) higher in SP. Five out of 89 data points (5.6%) are located outside of the LoA which showed a lack of agreement between the MS of both threshold perimetry in all the subjects using Bland-Altman plots. Further evaluation of the mean differences of MS using regression test showed no statistically significant proportional bias ($t = 1.180$, $p = 0.241$).

The Bland-Altman plot of MS in the glaucoma group only (n = 39) is shown in Figure 5.3. The bias/mean difference (95% LoA) between the strategies was reduced to 0.37 dB (-5.08, 5.82 dB). It was still lack of agreement shown in the glaucoma patients with 3 out of 39 data points (7.7%) located outside of LoA. No significant proportional bias was found in mean differences of MS using regression test ($t = 0.888$, $p = 0.381$)

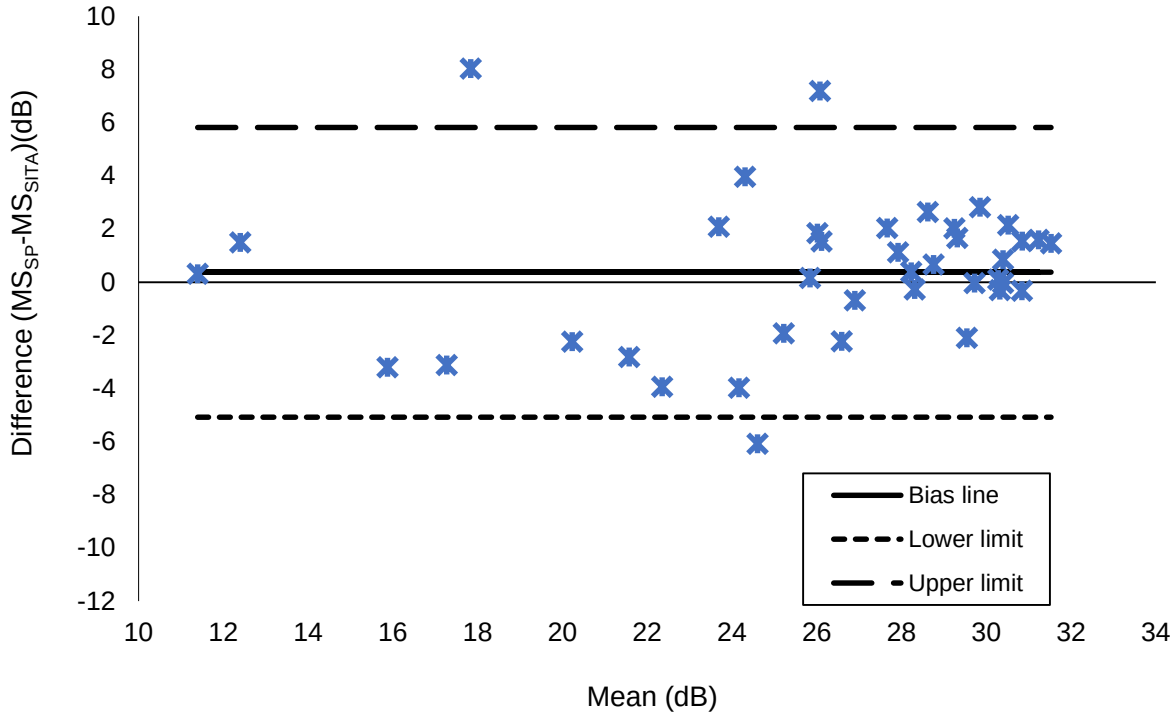


Figure 5.3: Bland-Altman plot of mean sensitivity between SS and SP in glaucoma patients

Bland-Altman plot also was used to determine the agreement between the strategies in the age-matched control group which is shown in Figure 5.4. The bias/mean difference (95% LoA) in MS of normal subjects between SS and SP was 0.46 dB (-1.44, 2.37 dB). There was 97.8% of the data points of normal subjects are located within LoA which indicates the agreement was achieved between the MS of the matching 66 test points of SS and SP in normal subjects (n = 45). There was also no statistical significant proportional bias on the differences (t = -0.527, p = 0.601) using regression test.

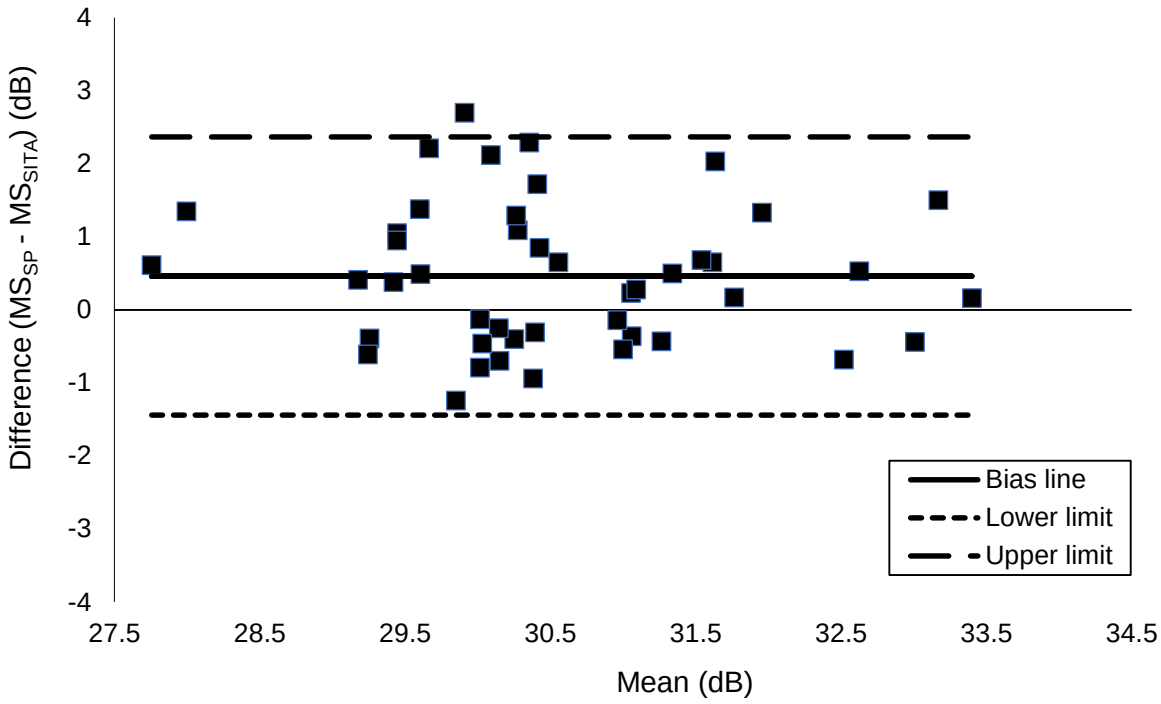


Figure 5.4: Bland-Altman plot of mean sensitivity using SS and SP in normal subjects

5.4.3.2 Mean deviation

The mean deviations were compared according to the recommendation by the manufacturer. Mean deviation (MD) in SS is the weighted average of the sensitivity deviations from the age-adjusted normal values for all the 76 test points used in 30-2 whereas MD in SP is non-weighted average from the 66 test points. A statistically significant correlation was found (Spearman correlation coefficient: rho = 0.885, $p < 0.001$) between the MD of SS and SP in all the subjects.

Figure 5.5 shows the Bland-Altman plot of MD between SS and SP in all subjects (n = 89). MD in SP was higher than SS with the bias/mean difference (95% LoA) between strategies of 2.44 dB (-1.12, 6.00 dB). Almost 9.0% of data points are located outside of LoA and statistically significant proportional bias on the difference between MDs of the threshold strategies (t = -3.235, p = 0.002) was shown. Lack of agreement between the MDs of both threshold strategies was indicated.

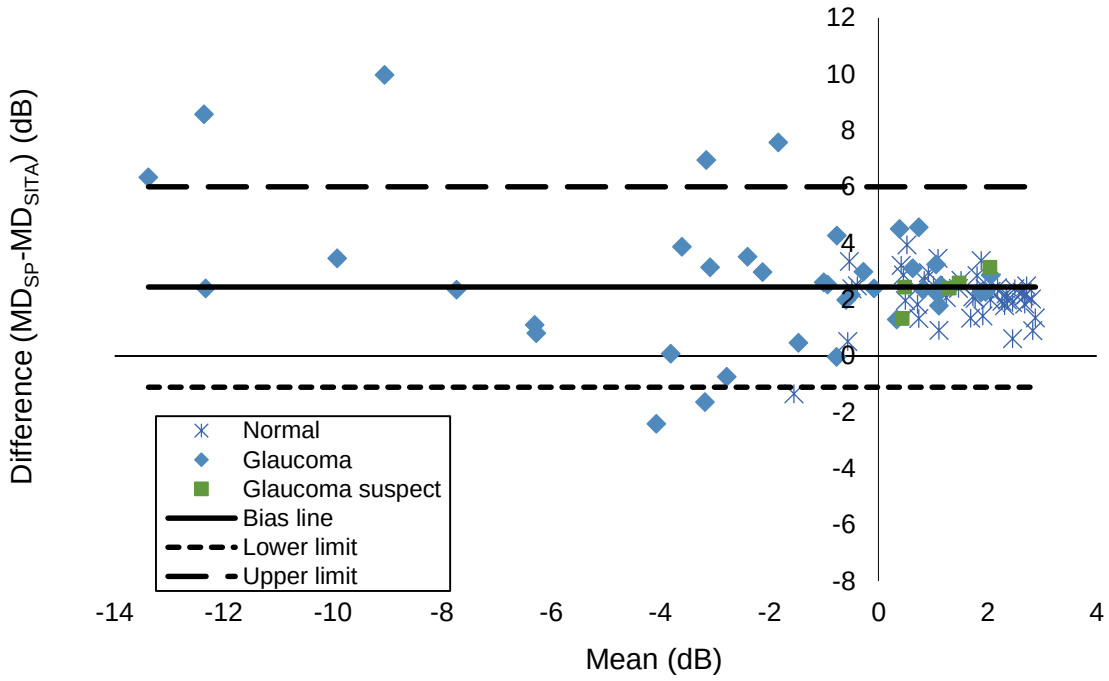


Figure 5.5: Bland-Altman plot of mean deviation between SS and SP in glaucoma, glaucoma suspects and normal subjects

Lack of agreement of MD between the strategies was also shown in glaucoma patients only using the Bland-Altman plot (Figure 5.6). There were 3 out of 39 data points located outside of LoA (7.7%) for glaucoma patients. A significant proportional bias was also found (t = -3.235, p = 0.02) using regression test. It showed a higher bias with lower MD value.

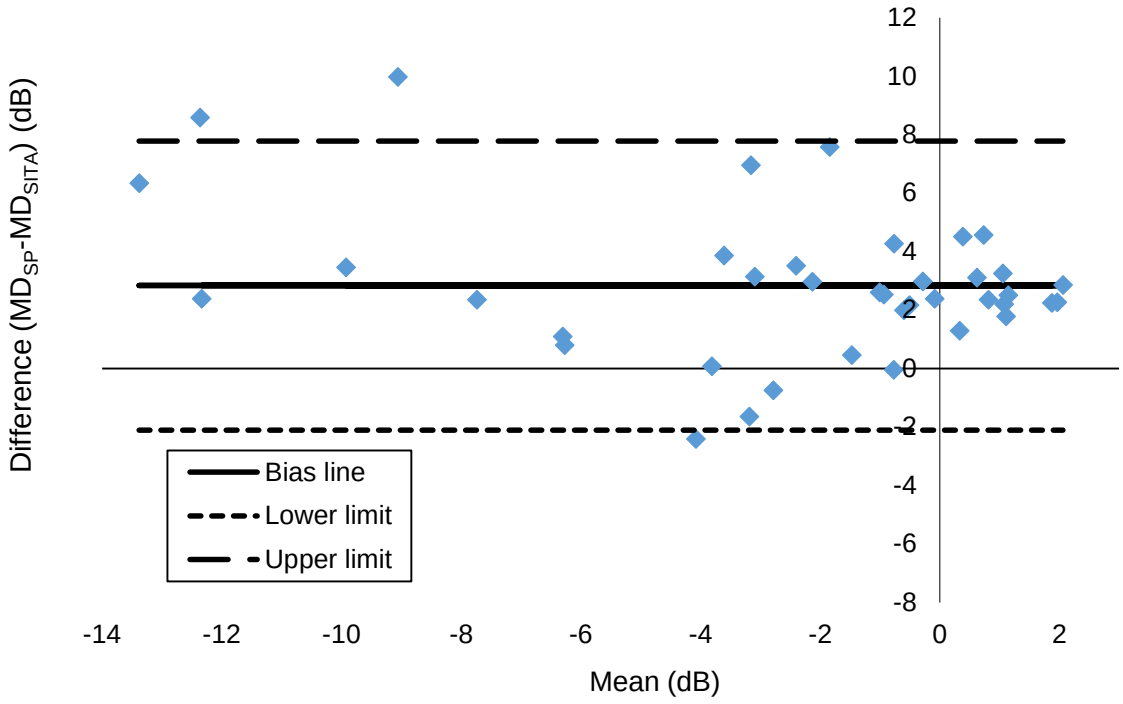


Figure 5.6: Bland-Altman plot of mean deviation between SS and SP in glaucoma patients

The Bland-Altman plot was also used for normal subjects only (Figure 5.7). The bias/mean difference (95% LoA) of MD between SS and SP was 2.84 dB (-2.11 dB, 7.78 dB) in glaucoma patients and 2.11 dB (0.35 dB, 3.87 dB) in normal subjects. Less than 5% of the data points (4.4%) are located outside of LoA for normal subjects and no statistically significant proportional bias (t =0.436, p = 0.665) was found. Thus, an agreement was found between the MD of SS and SP in normal subjects despite considerably large bias was observed.

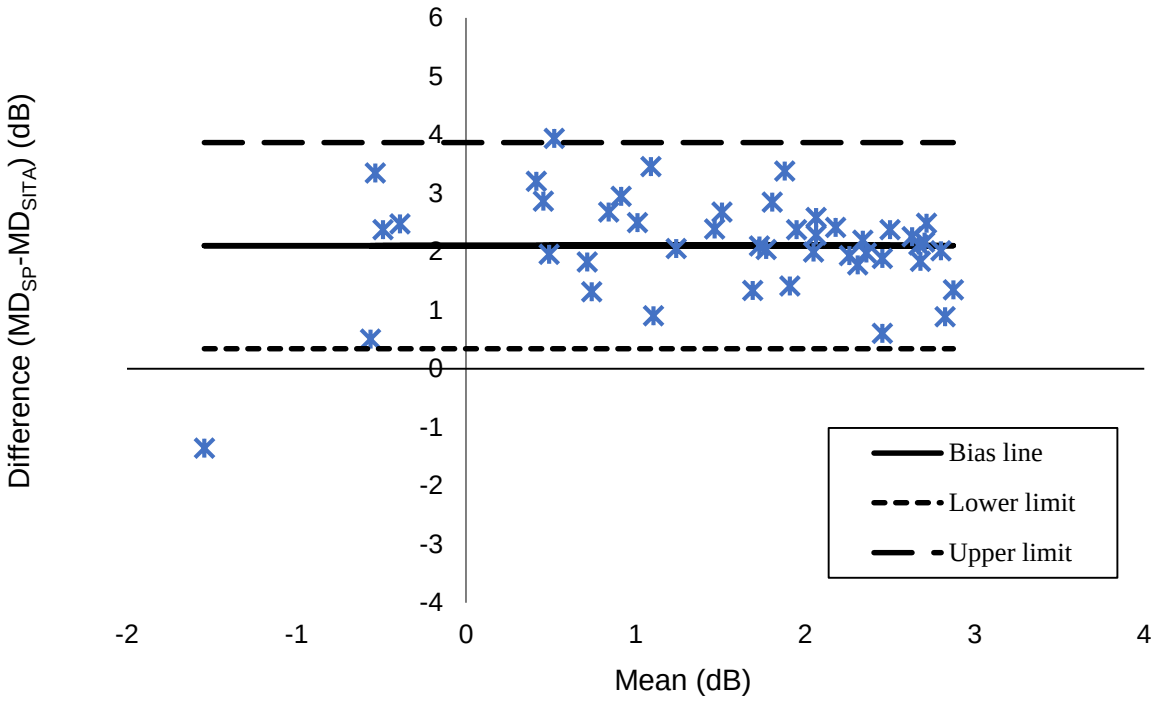


Figure 5.7: Bland-Altman plot of mean deviation between SS and SP in normal subjects

5.4.3.3 Pattern standard deviation

Pattern standard deviations between both strategies were shown to have statistically significant correlation (Spearman correlation coefficient: rho = 0.689, $p < 0.001$) in all the subjects (n = 89). Bland-Altman plot of PSD showed 9.0% of the data points located outside of LoA (Figure 5.8) which indicates the agreement was not achieved between the strategies for PSD. The bias/mean difference (95% LoA) between PSD of SS and SP was -1.52 dB (-5.55 dB, 2.50 dB) with higher PSD was found with SS. Further regression test on the differences showed statistical significant proportional bias for the differences of PSD (t =-14.390, $p < 0.001$). The differences between the PSD of the strategies were larger in higher PSD.

.

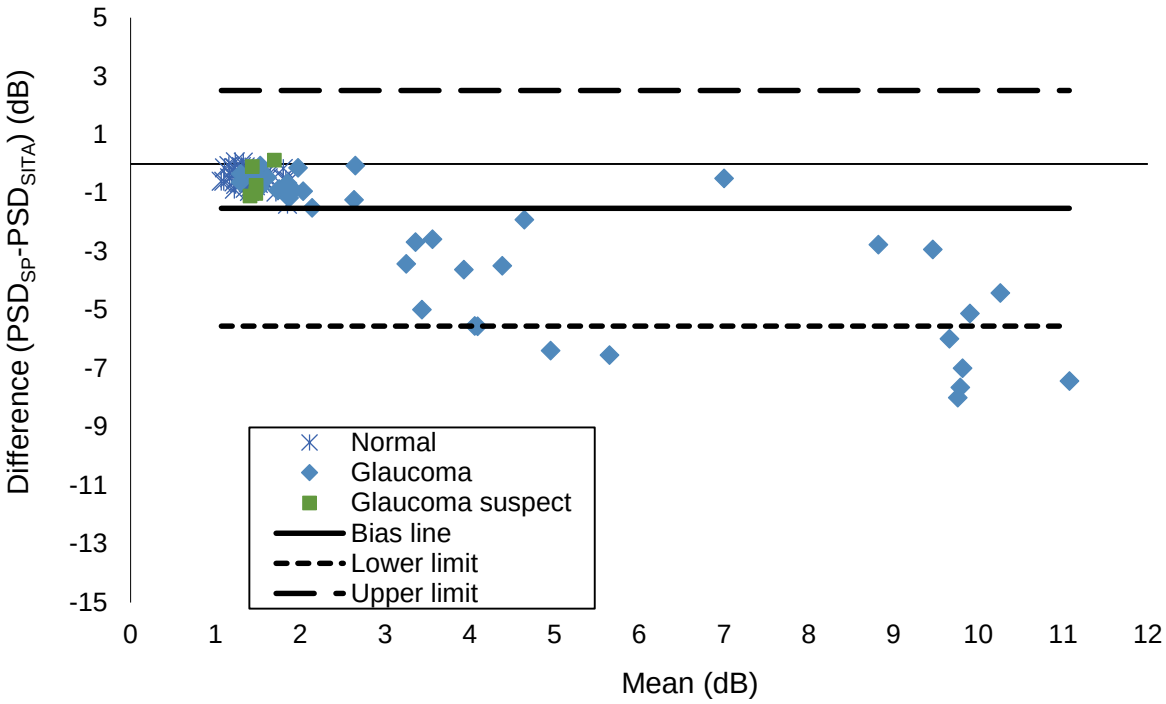


Figure 5.8: Bland-Altman plot of pattern standard deviation (PSD) between SS and SP in glaucoma patients, glaucoma suspects and age-matched normal subjects

.

Bland-Altman plot for the agreement between the PSD of SS and SP in glaucoma patients only showed 2.6% of data points located outside of the LoA with mean bias of -2.84 dB (-7.82, 2.15 dB) (Figure 5.9). Pattern standard deviation (PSD) from SS always showed a higher value than SP and an agreement could be achieved between the PSDs of the strategies according to the criteria with more than 95% of data points located within the LoA. However, large LoA of 9.97 dB indicates the bad agreement was obtained between the strategies. Furthermore, a regression test on the PSD differences showed statistically significant proportional bias ($t = -7.302$, $p < 0.001$) which indicates a larger difference between strategies in deeper field defect.

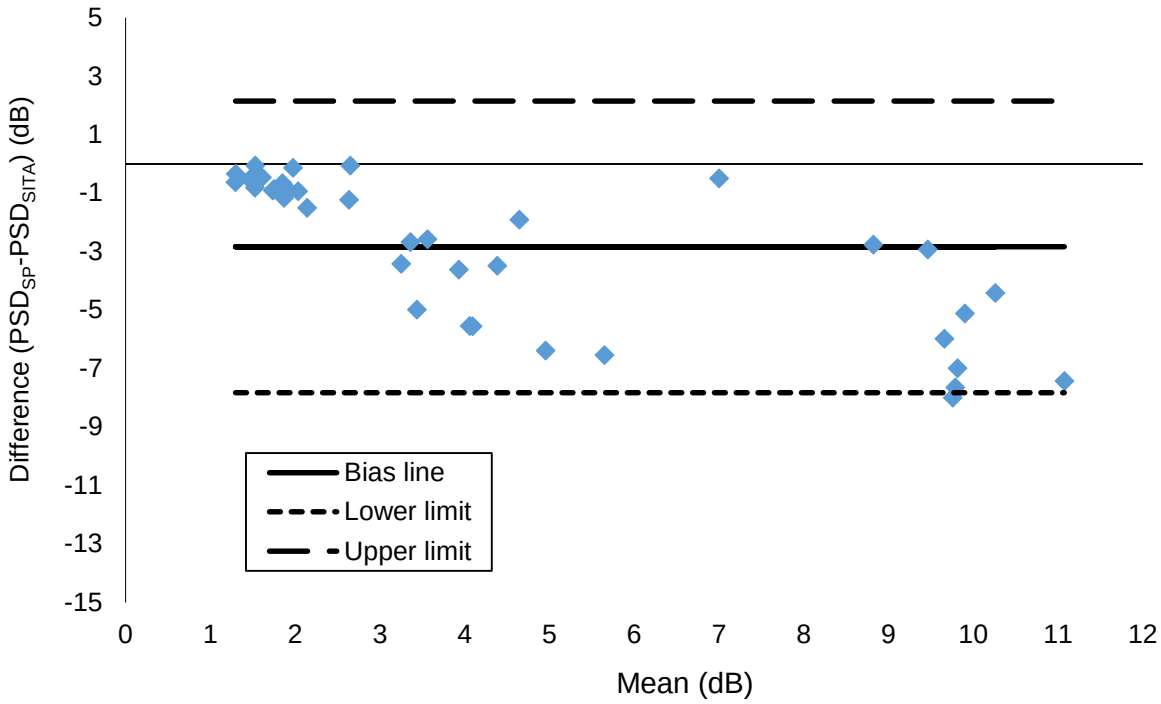


Figure 5.9: Bland-Altman plot of pattern standard deviation between SS and SP in glaucoma patients

Agreement between the PSD of the strategies was also achieved if the Bland-Altman plot analysis was used in normal subjects only (Figure 5.10). Only 2.2% of the data points are located outside of the LoA with bias/mean difference (95% LoA) of PSD between SS and SP was -0.49 dB (-1.05, 0.08 dB). However, there was also a statistically significant proportional bias for the differences of PSD (t = -2.224, p = 0.031).

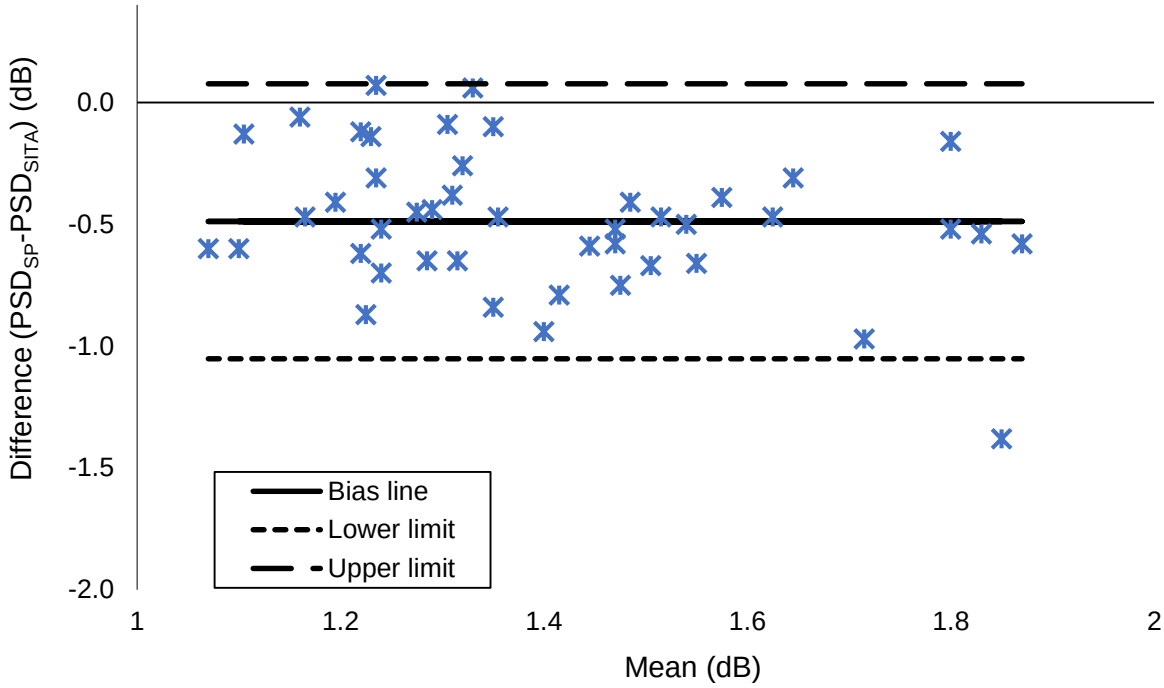


Figure 5.10: Bland-Altman plot of pattern standard deviation between SS and SP in normal subjects

5.4.4 Pointwise Between-strategy Analysis

Pointwise between-strategy comparison using Bland-Altman agreement analysis which mean differences and 95% LoAs of sensitivities were determined at each 66 test points as shown in Figure 5.11. There was no apparent association between the bias and eccentricity of the stimulus locations as well as apparent inter-hemifield difference between superior and inferior field. Nevertheless, there was a marked difference of bias shown between temporal and nasal field. It showed that the SP produced higher threshold estimates than SS mostly in the nasal field whereas higher threshold estimates from SS were found more in the temporal field.

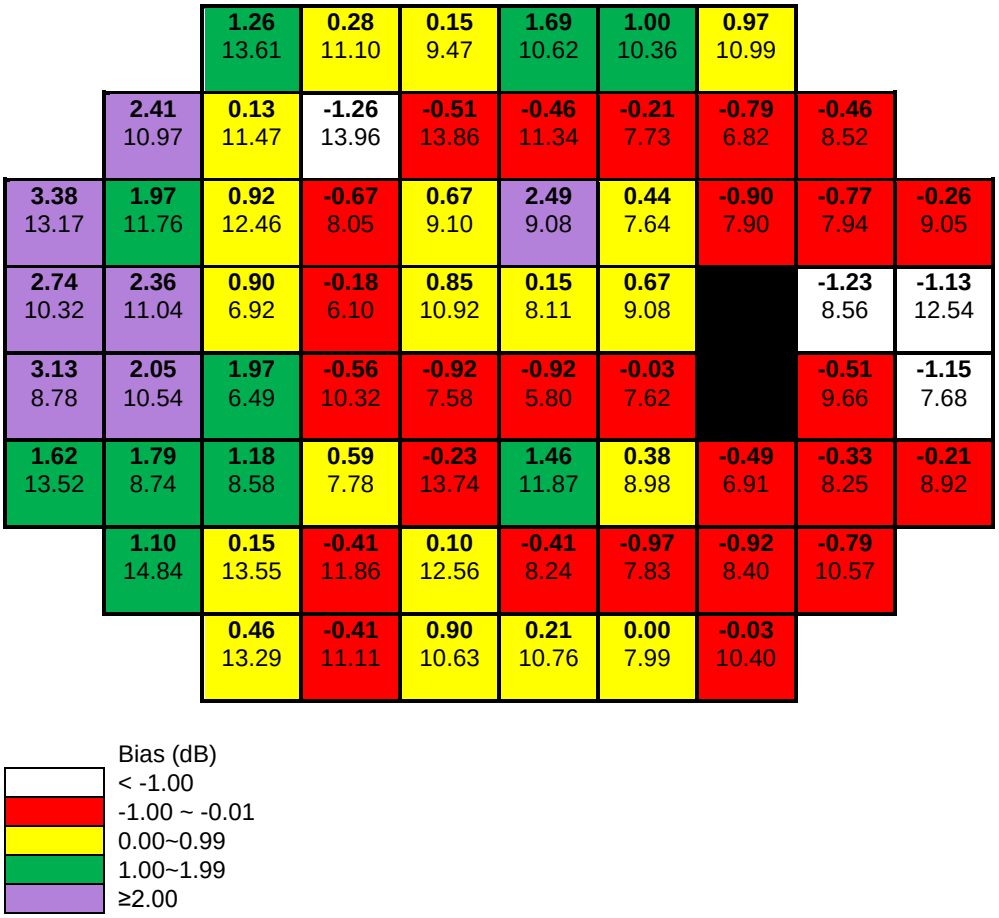


Figure 5.11: Bland-Altman agreement analysis of MS between SS and SP at the 66 test points in glaucoma patients. Bias/mean difference (upper and bolded number) and the 95% limits of agreement (lower and unbolded number) in the threshold estimates for each 66 matching test points between SS and SP.

*The bias /mean difference is the MS of SP minus that of SS
*The colours represent the ranges of the bias/mean difference in MS between SS and SP

5.4.5 Between-strategy Comparison of AGIS Score

AGIS score was calculated for each glaucoma patients based on SS and SP results. The data of the 39 AGIS scores calculated were not normally distributed according to the Shapiro-Wilk test (see Appendix A5.3). The comparison of AGIS scores between SS and SP is shown in Table 5.6.

Table: 5.6 AGIS score using SITA Standard and SPARK Precision in glaucoma patients

	AGIS score				
	Mean	SD	Median	Min	Max
SS	3.69	4.76	1.00	0	14
SP	1.96	3.23	0.00	0	10

There was a statistically significant difference between AGIS scores of SS and SP (Wilcoxon Signed Rank test: $Z = -3.767$, $p < 0.001$). Nevertheless, the AGIS scores from both strategies were statistically significant positively correlated (Spearman correlation coefficient: $rho = -0.750$, $p < 0.001$).

There are recommended glaucoma stages according to AGIS score (The AGIS, 1994) i.e. 1 to 5 mild, 6 to 11 moderate, 12 to 17 severe, 18 to 20 end-stage of glaucoma. The distribution of the stages in both SS and SP is shown in Table 5.7. More subjects were categorized in less severe stage according to SP results in comparison to SS.

Table 5.7: Distribution of the glaucoma patients according to the glaucoma stages based on AGIS score using SS and SP

Glaucoma stage	Total of glaucoma patients			
	SITA Standard		SPARK Precision	
	N	%	N	%
None	16	41.0	26	66.7
Mild	12	30.8	5	12.8
Moderate	6	15.4	8	20.5
Severe	5	12.8	0	0.0
End-stage	0	0.0	0	0.0

5.4.6 Size and depth of the glaucomatous field defect

The size and depth of the glaucomatous field defect according to the number and sensitivity deviations of the abnormal pattern deviation points based on HPA criteria (Hodapp et al., 1993) were calculated and compared between the two strategies as shown in Table 5.8.

Table 5.8: The size and depth of the glaucomatous field defect according to the results of SS and SP

	Defect size (dB)					Defect depth (dB)				
	Mean	SD	Median	Min	Max	Mean	SD	Median	Min	Max
SS	9.0	7.6	8	0	26	131.9	159.7	48	0	571
SP	6.7	10.9	0	0	24	80.6	135.9	0	0	371

The data from both strategies were not normally distributed according to the Shapiro-Wilk test (see Appendix A5.4). No statistically significant difference between defect size of SS and SP (Wilcoxon Signed Rank test: $Z = -1.511$, $p = 0.131$) but statistically significant deeper field defect was displayed by SS than SP (Wilcoxon Signed Rank test: $Z = -3.048$, $p = 0.002$). Either size or depth of glaucomatous field defect detected by the strategies showed a statistically significant positive correlation between the two strategies (Spearman correlation coefficient: rho =0.589, $p < 0.001$ for defect size; rho = 0.763, $p < 0.001$ for defect depth).

The ratio between the total defect depth and total defect size which is the average depth for each abnormal stimulus point was determined for each strategy (Table 5.9). SITA Standard clearly produced the results with more severe localized field defect. Figure 5.12 shows an example of the comparison of the glaucomatous VF results between SS and SP which show deeper field defect per single abnormal stimulus point when using SS according to HPA criteria (Hodapp et al., 1993). There were 14 eyes showed glaucomatous field defect with SS but appeared to be normal with SP. One of the examples is shown in Figure 5.13.

Table 5.9: Average depth of each abnormal stimulus point for SS and SP

	Defect size (dB) Total	Defect depth (dB) Total	Average (Depth/Size)
SS	351	4325	12.3
SP	261	2496	9.6

Threshold Graytone SPARK Precision

Pattern Deviation Analysis Map

MD: -5.75 dB PSD: 7.36 dB

SITA Standard 30-2

MD: -6.85 dB PSD: 14.79 dB

Figure 5.12: Example of glaucomatous VF results for the comparison between SPARK Precision and SITA Standard which show the area of the field defect according to HPA criteria (Hodapp et al., 1993). By referring to the inferonasal glaucomatous field defect outlined in red, the ratios of depth/size for SP and SS are 13.1 dB and 21.5 dB respectively

Threshold Graytone Pattern Deviation Analysis Map
SPARK Precision

MD: -0.64 dB PSD: 0.94 dB

SITA Standard 30-2

MD: -4.16 dB PSD: 5.93 dB

Figure 5.13: The upper VF result is normal with SPARK Precision. The lower result from the same eye showed glaucomatous VF results by SITA Standard. The area of glaucomatous field defect according to HPA criteria (Hodapp et al., 1993) is only detected in SS which is outlined in red. The ratios of depth/size for SP and SS are 0 dB and 10.6 dB respectively

5.4.7 Sensitivity and Specificity in Detecting Glaucomatous Field Defect

The sensitivity and specificity of SP in detecting the glaucomatous field defect in 89 subjects (39 glaucomas, five glaucoma suspects and 45 normals) using the results of SS as reference standard shown in Table 5.10. The criteria of detecting glaucomatous field defect were according to HPA criteria (Hodapp et al., 1993) for SS and Gonzalez de la Rosa et al. (2013) and manufacturer's recommendation for SP. The present study had shown the sensitivity of

using SP to detect glaucomatous field defect as judged by the results of SS was merely 57.5% but there was not a single false positive result from SP which contributed to 100% of specificity.

All the glaucoma cases were confirmed the diagnosis by local ophthalmologists before the treatment is initiated. The sensitivity of using SS in detecting the glaucomatous eye was 84.6% whereas SP was only able to detect 48.7% of the glaucomatous eye. The sensitivity and specificity indices of both strategies according to each criterion stated earlier against the diagnosis made by local ophthalmologists were determined as shown in Table 5.11. According to each criterion, the sensitivity of SS to diagnose glaucoma was as high as 84.6% whereas SP was only able to achieve up to 48.7% of sensitivity. Among the criteria using SP, the criterion using the PSD was the most useful criterion in detecting glaucomatous eye whereas the criterion using MD was the least sensitive.

Table 5.10: Glaucomatous field defect detected by SP vs SS

	Visual field defect according to SS		
	Present	Absent	Total
SP positive	19	0	19
SP negative	14	56	70
Total	33	56	89

SP sensitivity = 19/33 × 100 = 57.5%
SP specificity = 56/56 × 100 = 100%

Table 5.11: Sensitivity and specificity of SS and SP in detecting glaucoma according to each of the criteria against the diagnosis by local ophthalmologists

	SITA Standard			SPARK		
Criterion	SS(a)	SS(b)	SS(c)	SP(a)	SP(b)	SP(c)
Sensitivity	84.6%	59.0%	61.5%	30.8%	48.7%	33.3%
Specificity	100%	100%	100%	100%	100%	100%

SS(a) = A cluster of three or more non-edge points in either hemifield on pattern deviation probability map with sensitivity found at p<5% with at least one of the defective points having p<1%;
SS(b) = PSD had a value at p<5%;
SS(c) = GHT showed "outside normal limits".
SP(a) = MD < -2.3 dB
SP(b) = PSD > 1.8 dB
SP(c) = More than 5 defect points/scotomas with total deviation > 5dB.

The ability of SP and SS to detect VF defect in different severity level of glaucoma according to the AGIS score was determined (Table 5.12). It showed poor sensitivity of SP with only 13% sensitivity in detecting eyes with mild glaucoma defects as compared to using SS which achieved 69% of sensitivity.

Table 5.12: Sensitivity of SS and SP in detecting glaucomatous field defect according to the glaucoma severity level based on AGIS score

AGIS score	Category	No of patient	Sensitivity	
			SS	SP
0	None	16	69%	13%
1 - 5	Mild	12	100%	58%
6 - 11	Moderate	6	100%	83%
12 - 17	Severe	5	100%	100%

5.5 Discussion

This study examined the performance of SP compared to the well-recognized gold standard threshold strategy, SS in assessing glaucoma. To date, there is no data published comparing these two threshold strategies even though SPARK is commercially available since 2011. SPARK is incorporated in Oculus Perimeter which is not widely used in Malaysia but it is widely used in other countries such as Germany and Spain. The comparison of SPARK with the current “gold standard” and well-accepted threshold strategy, SITA is important for SPARK to become a clinical alternative in the threshold perimetry particularly for diagnosis and monitoring of glaucoma with the advantage of its evidently lower testing time.

In this study, distinct differences were observed between the global indices of SP and SS particularly MD and PSD either in glaucoma or normal subjects. Mean deviations between the strategies were highly correlated but the agreement between the MD of the strategies was only achieved in normal subjects. The bias of MD between-strategy was as high as 2.84 dB in glaucoma subjects and 2.11 dB in normal subjects which were considerably overestimated by SP. The bias even increased with poorer sensitivity in glaucoma subjects. Mean deviation was regarded as the index from SPARK that provides maximum diagnostic capacity for glaucomatous defect (Gonzalez de la Rosa et al., 2013). Overestimated MD could affect the sensitivity in diagnosing glaucoma which was also shown in this study as well. The larger differences between the MD of the strategies as compared to the difference between MS could be attributed to the following factors:

a) Different methods of MD calculation are applied by SITA and SPARK. The MD of SITA is weighted according to eccentricity i.e. the deviation from the age-matched normal value at each point is divided by the variance of normal values at that point (Funkhouser and Fankhauser, 1991). As the variance increases with greater eccentricity, peripheral test points are less important in the calculation. Conversely, the average of the deviation in SPARK is calculated without the weighting procedure. Nevertheless, Funkhouser and Fankhouser (1991) claimed the difference between the weighted and

non-weighted indices was negligible and both indices are interchangeable whereas Heijl et al. (1992) argued that the weighting method improves the accuracy of the threshold estimation. It also provides an advantage in the perimetric analysis which helps in the detection of glaucoma (Asman and Heijl, 1992b). Flanagan et al. (1993b) investigated the effect of the weighting factor on global indices and found no influence on MD but slight increment of PSD. Nevertheless, the weighting method used in the SS possibly contributed to some differences in global indices compared to SP.

b) MD from SS 30-2 is calculated according to the 76 test points whereas MD from SP is calculated from the 66 test points including the central threshold but without the uppermost and bottommost rows of test points in 30-2 which could be easily affected by the artefact of drooping eyelid and correcting lens (Wang and Henson, 2013). Recalculation of MD by using the same number of test points was not performed in the same way as the calculation of MS in this study considering the recalculation is not practical in the clinical practice and the values of the global indices provided in the printout of the result are more commonly used by practitioners for interpretation. On top of that, as peripheral test points are weighted less to the calculation of MD, the impact of this factor to the difference of MD between the strategies could be low.

c) Different age-matched normative databases are used in the strategies. The details of the database for the age-corrected sensitivities used in the strategies have never been published. The database could provide a depressed normal hill of vision which resulted in more positive threshold deviation test points and ultimately higher MD and poorer PSD. The influence of the database to the differences between-strategy need to be further investigated and it is believed to have played an important role in the differences observed in this study.

The average MD from the group of normal subjects in this study was 2.58 dB (SD 1.20 dB). It was apparently higher than the values reported by Gonzalez de la Rosa et al. (2013) which were only 0.13 dB to 0.23 dB using Easyfield perimeter. This could be partly due to the different

age range of the subjects recruited which was not reported by Gonzalez de la Rosa et al. (2013) but according to the regression rate of threshold sensitivity against age in Chapter 4 (-0.076 dB/yr), it needs more than 20-year difference in mean age between groups of normal subjects to justify the difference. Therefore, higher bias due to the difference in the mean age of the groups is unlikely. Differences in instrumentation and normative database are likely the causes of the bias. Further rectification is required to identify the reason for the higher bias.

The difference in the PSD between strategies is also affected by the same reasons for the differences between MDs mentioned earlier. Moreover, the effect of the short-term fluctuation is also taken into account for the calculation of PSD in SITA (King et al., 2002). In this study, PSD was always higher in SS than in SP indicated deeper local field defect (Wall and Johnson, 2005) was measured by SS. The agreement between-strategy for PSD could not be achieved when Bland-Altman analysis was carried out for all the subjects or glaucoma patients only. Large LoA between strategies was found in the glaucoma group. It showed the differences between strategies had become larger with more advanced glaucomatous field defect. Even though the agreement between-strategy was achieved in normal subjects, it showed the tendency of the difference to be increased when lower sensitivity was found. The normal subjects that showed much smaller differences in PSD between the strategies as compared to glaucoma subjects mainly due to decreased irregularities within the normal field than the glaucomatous field (Yaqub, 2012). The mean value of PSD in normal subjects reported in this study was comparable to the value reported by Gonzalez de la Rosa et al. (2013) but relatively lower PSD values were reported in glaucoma group recruited in the current study.

The size of the glaucomatous field defect detected by SP was not significant different compared to the one detected by SS but the depth was shallower when using SP. It is questionable that which strategy would produce the “real” threshold sensitivity. SPARK Precision showed shallower defects which could be due to the average value obtained from the repeated tests from the 4 phases. The one with the most extreme value is excluded

(Gonzalez de la Rosa et al., 2013). It is a good way to reduce the variability of the threshold estimates but it may also mask the field defects which ultimately underestimate the severity of the disease. It is also associated with the lower value of PSD produced by SP which indicates a shallower localized defect.

SPARK Precision had shown apparently low sensitivity in detecting glaucomatous eye either against the results using SS or diagnosis of local ophthalmologist that relied on the combination of functional and structural examination. It was attributed to the shallower localized defect detected by SP and the number of mild field defects were undetected. The criteria used to diagnose glaucomatous field defect depend on the values of global indices and deviation map which were shown to be substantially and significantly lower when using SP. There were about 14 out of 33 patients with glaucomatous field defects detected by SS but undetected by SP and about half of the glaucoma patients diagnosed by local ophthalmologists produced normal VF results with SP whereas SS only missed about 15% of them.

The poor diagnostic capability of SP had become more apparent in the early stage of glaucoma which showed only 13% of sensitivity as compared to SS (69%) in detecting the glaucomatous field loss. The use of the SP especially for the diagnosis of the early glaucoma is definitely not recommended according to this study.

With regard to SS, its sensitivity in detecting glaucoma eye in this study achieved nearly 85% which is reasonable according to the previous studies that were ranged from 83% to 98% depending on different criteria and severity level of the glaucoma (Budenz et al., 2002a; Sekhar et al., 2000; Sharma et al., 2000; Wadood et al., 2002). The specificity of the SS reported in this study was not reliable as the normal eyes showed VF defect with SS were rejected. It was only used as a reference to determine the specificity of SP which was shown 100% against the results from SS. None of the normal eyes had shown significant field defect with SP.

Gonzalez de la Rosa et al. (2013) recommended to use MD < -2.3 dB and PSD > 1.8 dB in the result of SPARK as the cutoff point for the diagnosis of VF defect which achieved the sensitivity of 86.5% and 82.7% respectively in their study. However, in the current study, substantially poorer sensitivities of less than 50% were reported using the same criteria. The third criterion used in SPARK was according to the number of the defect points that each has a total deviation of more than 5 dB. Poorer sensitivity of only 33.3% was shown as compared to 80.9% as reported by Gonzalez de la Rosa et al. (2013) with 5 defect points was used as the cut off value. Despite using the same SPARK strategy in both studies, perimeters used in the studies were different. Easyfield perimeter (Oculus Optikgeräte GmbH, Wetzlar, Germany) was used in the study of Gonzalez de la Rosa et al. (2013) whereas Oculus Twinfield 2 (Oculus Optikgeräte GmbH, Wetzlar, Germany) was used for the current study. Both perimeters have similar essential parameters including background illumination, stimulus size, stimulus duration and interval. The maximum stimulus luminance in Twinfield is achieved through simulated method whereas the Easyfield perimeter has its maximum stimulus luminance up to 10,000 asb. Physically, Twinfield 2 could only display stimulus luminance maximum up to 318 cd/m^2 (1000 asb). Without changing the maximum luminance in the perimeter, the calculation method of the decibel scale is modified instead. The decibel scale is used to describe the difference of magnitude of luminance relative to a specified reference level which is usually the maximum stimulus luminance that the perimeter can produce (Refer Eq. 5.1)

$$\text{Decibel} = 10\log (L_{max}/L) \qquad \text{(Eq. 5.1)}$$

L_{max} – the maximum stimulus luminance that the perimeter can display (reference value for decibel scale)

L – the luminance of the stimulus at the threshold

(Racette et al., 2016)

The simulated maximum stimulus luminance in Twinfield 2 is achieved simply by modifying the reference value in order to display decibel scale that corresponds to the same luminance values as in Humphrey or Easyfield perimeter (10,000 asb). Moreover, if the difference in the

result was due to the difference in instrument set-up, a marked difference should also be exhibited between the MSs which otherwise would show only a small difference in this study. The age-corrected normal database used in SPARK needs to be further verified when conducted using Twinfield perimeter to identify the possible cause for the overestimation of MD and underestimation of PSD which eventually compromise the diagnostic capability of the SP. Capris et al. (2008) had also pointed out the doubt in the normal database used in Oculus CLIP which resulted in an underestimated MD.

AGIS scoring system is an objective quantitative scoring system that use HFA 24-2 total deviation map to produce a single index score (The AGIS investigators, 1994). The use of 30-2 in SS and 30 x 24 test area in SP for the current study did not affect the calculation of the AGIS score. In this study, the AGIS scores obtained from both strategies were correlated, this was reasonable as the global indices were proven to be correlated between the strategies. Nevertheless, there was a significantly lower AGIS score calculated using the results from SP and many cases skewed to lower level of severity compared to the score using SS. It was mainly due to the AGIS scoring system was referring to the total deviation map which obviously had higher MD with SP.

The development of SPARK was based on the hypothesis that using a minimal number of test points could still suffice to detect localized defects without compromising the diagnostic capability (Gonzalez de la Rosa et al., 2013). By considering these presented results, the aforesaid hypothesis is questionable. Overestimation of MD had resulted in less diagnostic sensitivity when using the recommended criteria but was shown otherwise in the study by Gonzalez de la Rosa et al. (2013). The same for the underestimated PSD even though it provided better diagnostic capability than using MD in this study (see pg. 159)

Apart from the results of MD and PSD, the MSs between the strategies were actually not markedly different. Even though the present study showed a statistically significant higher MS

with SP compared to using SS in control subjects, but the mean difference merely reached 0.5 dB. It was unlike to be clinically significant for a routine clinical assessment of VF test. Similar finding was shown in the previous chapter (see Chapter 4, p. 119) that higher values in SP were partly due to reduced fatigue effect (Gonzalez de la Rosa and Pareja, 1997; Hudson et al., 1994; Heijl and Drance, 1983) with more than 40% reduction of testing time when using SP. Such MS differences between strategies were not statistically significant in glaucoma patients even though more than 50% of the testing time was reduced and the magnitude of the differences was larger. It was partly ascribed to the larger inter-subject variability within the glaucoma group with the range of threshold estimates recorded almost four times larger than the normal group. Significance level required was broader which was also attributed to the range of patient recruited consisted of early to severe stages of glaucoma in this study. The agreement between the MSs of the two strategies was also not achieved in glaucoma patients even though they were correlated and smaller bias was observed. The drawbacks were the high LoA and lack of data points within 95% LoA which were not helped by large inter-subject variability displayed by glaucoma patients. Greater dispersion of the data was found for threshold sensitivities below than 28 dB but the dispersion did not increase with depressing mean of threshold sensitivities. However, a good agreement of MS between the two strategies was showed in the control group with lower bias and LoA than those found in Chapter 4.

In this study, both strategies had produced significant lower threshold estimates in glaucoma patients who required significant longer testing time depending on the severity of the VF defect (Wild et al., 1999a; Roggen et al., 2001). SPARK Precision was able to produce more consistent testing time even in glaucoma patients with VF defects who had a longer test duration of only an average of six seconds compared to normal subjects. Based on the current study, the average testing duration using SS (6.3±0.7 min) was comparable to testing duration reported in Budenz et al. (2002) (6.6±0.7 min) for the normal subjects with an approximately similar mean age (51.9 vs 52.9). However, relatively longer testing time was required by glaucoma group in Budenz et al. (2002) as compared to the current study (8.8±1.4 min vs

7.7±1.6 min) which might be due to more severe glaucoma cases in the study published by Budenz et al. (2002). The present study was intended to recruit more early-stage glaucoma cases to assess the diagnostic capability of SPARK.

Further pointwise evaluation on the threshold estimates in the 66 test points shows the highest MS overestimated test points by SP were located at the nasal step i.e. the common area of VF defect in glaucoma (Risse et al., 1999; Hart and Becker, 1982; Lee et al., 2003). On the other hand, the most underestimated test points by SP were located at the temporal test points adjacent to the blind spot which are less likely to display glaucomatous field loss. An approximately similar pattern of the differences was also observed in the comparison between SS and Zippy estimation by sequential testing (ZEST) (Rao et al., 2017) which is a maximum-likelihood threshold determination algorithm used in Compass. Zippy estimation by sequential testing overestimated threshold sensitivity in the nasal field whereas SS produced lower values in temporal and central 15° field. The actual reason for the pattern is yet to be explained and Rao et al. (2017) had concluded that Compass cannot be used to replace HFA in the context of glaucoma management.

5.6 Conclusion

SPARK Precision overestimated MD as high as 2.84 dB of bias compared to SS in both glaucoma patients and normal subjects with higher bias found in the glaucoma patients. Agreement between MDs of both strategies could not be established in glaucoma patients but it was otherwise in normal subjects. The high bias of MD between the strategies may be related to the sceptical age-corrected normal database used in the total deviation analysis map which provides depressed normal hill of vision. Further evaluation of the normative database used is required. Conversely, an underestimated PSD was shown in SP as compared to PSD from SS. The agreement of PSDs between strategies was also not accepted in the glaucoma group but it was otherwise in normal subjects. There was a tendency of the bias increment with the

deeper localized defect and higher between-strategy bias was distinctly shown in glaucoma patients than in normal subjects.

The underestimated of the glaucomatous field defects by using SP had significantly reduced the diagnostic capability of the SP in detecting glaucoma. Poor sensitivity, especially in the diagnosis of the early stage of glaucoma eyes, was shown when using the results of SP. Nonetheless, it maintained high specificity with the results of SS as reference.

The use of the SP for the diagnosis of the early stage of glaucomatous field defect is not recommended unless further improvement and rectification on the age-corrected normative database used in the total deviation analysis is done.

Due to the overestimated MD with SP, AGIS scores based on the results of SP were lower than using the results of SS and it indicates the results of SP underestimate the severity level of the glaucomatous defect. The size of the glaucomatous field defect detected by SP was not significantly different with the size based on the results from SS but shallower depth of the defect was produced by SP.

Comparison between SS and SP showed more than 50% of the testing time was saved in glaucoma patients when SP was used but there was no significant difference of MS found between the strategies. It had also shortened the testing time of normal subjects by 40% which also showed significantly higher MS perhaps due to the reduced fatigue effect. The agreement of MS between strategies was only achieved in normal subjects despite low bias was also found in the glaucoma group. The greater difference between the MS of the strategies was observed for sensitivities below 28 dB.

Pointwise comparison between the threshold sensitivities of SS and SP showed SP tends to overestimate the sensitivity for the points located at nasal step while underestimating the points

located at the temporal field adjacent to the blind spot. The difference between the strategies was hypothesized to be associated with the common areas of glaucomatous field defect.

CHAPTER 6

COMPARISON BETWEEN SPARK PRECISION AND SITA STANDARD IN SUBJECTS WITH CATARACT

6.1 Introduction

A cataract is characterized by the opacity of the crystalline lens which is associated with old age (Na et al., 2014; Nowak and Smigielski, 2015; Asbell et al., 2005). It has remained as one of the leading causes of preventable blindness worldwide (Khairallah et al., 2015; Pascolini and Mariotti, 2012; Li et al., 2013) despite the cataract surgery is considered easily accessible, relatively safe and cost-effectively treatable (Lansingh et al., 2007; Lundstrom et al., 2015). The rapid growth of the ageing population could be one of the reasons why the growth rate is higher than the ophthalmologists' (Rao et al., 2011). Due to the prevalence of glaucoma is also higher in the aging population (Bourne et al., 2016; Kapetanakis et al., 2016) coupled with a higher risk of cataract after glaucoma treatment (The AGIS investigators, 2001; Heijl et al., 2002), cataract could be found frequently co-exist with glaucoma. The effect of cataract on the VF examination has been well-established (Wood et al., 1989; Klein et al., 1996). In threshold perimetry, cataract was reported to cause general reduction of threshold sensitivity (Lam et al., 1991; Budenz et al. 1993; Rehman Siddiqui et al., 2007; Koucheki et al., 2004; Smith et al., 1997; Hayashi et al., 2001; Kook et al., 2004; Musch et al., 2006; Ang et al., 2010; Kim et al., 2001). It is important to recognise and isolate the effects of cataract and glaucoma on threshold estimation. Thus, the progression of glaucoma (If co-existent) can be masked in the VF test by the co-existence of the cataract. The effect of cataract on threshold perimetry was already reviewed in Chapter 1.7 (pg. 67). A clinical valid VF test should be able to differentiate optical or neural field loss and detect any field loss in the early stage. Cataract causes typical optical origin field namely diffuse loss whereas glaucoma is the common contributor of neural field defect manifesting as a focal loss. In order to differentiate between diffuse and focal loss, cataract patients (free from glaucoma) were recruited for this study to assess the capability

and reliability of SPARK Precision (SP) compared to SITA Standard (SS) in the estimation of sensitivity value.

Swedish Interactive Threshold Algorithm (SITA) has been the gold standard of threshold strategy widely used in numerous studies not just to assess the glaucomatous eyes but also for eyes with cataract (Chung et al., 2016; Rehman Siddiqui et al., 2007; Rao et al., 2013; Ang et al., 2010). The global indices, total deviation (TD) and pattern deviation (PD) map are common analysis tools used to segregate the effect of cataract and neural defect on the VF result (Bengtsson et al., 1997a; Katz, 2000; Rehman Siddiqui et al., 2007). Oculus has introduced SPARK for the similar field analysis tools which are capable to identify the effect of cataract on the VF. The advantage of SPARK is 50% shorter testing time in glaucoma patients compared to SS (Refer Chapter 5, pg. 144) but overestimates MD and underestimates PSD when compared to SITA. This can lead to a lack of detecting VF focal loss which is the typical sign in glaucoma patients (Refer Figure 5.13, pg. 159). Its global indices i.e. MS, MD, and PSD did not agree with indices from SITA in a group of subjects comprised of glaucoma, glaucoma suspect and normal subjects. Hence, this study will focus on the evaluation of the performance of SP in detecting the eyes which commonly exhibit more diffuse sensitivity loss i.e. cataract patients (Budenz et al., 1993; Lam et al., 1991; Guthauser and Flammer 1988; Wood et al., 1989). In comparison to the established SITA, it could provide a more complete understanding of the relatively new threshold strategy.

6.2 Objectives

Following from the last two chapters with the threshold strategies of SPARK Precision (SP) using Oculus Twinfield 2 and SITA Standard (SS) using Humphrey Field Analyzer (HFA) were compared in healthy normal subjects and glaucoma patients, the performance of SP in estimating the threshold sensitivities was further evaluated against current gold standard threshold strategy, SS in subjects with cataract (exhibiting diffuse VF loss). The main objectives of this study are as follows:

a) To determine and compare the extent of the global differences of the threshold sensitivities from the matching 66 test points between SP and SS in subjects with cataract and also in age-matched healthy control subjects
b) To compare the global indices (i.e. MD and PSD) and the time duration between the strategies in cataract and control subjects
c) To determine the correlation of the global indices between the strategies in all the subjects including cataract and control subjects
d) To determine whether the agreement is established between the strategies for MS, MD, and PSD in subjects with cataract or control subjects and/or both
e) To evaluate the extent of the pointwise differences of the threshold estimates between the two strategies in all 66 test locations in subjects with cataract

6.3 Methods

6.3.1 Research Participants

The cataract and normal control subjects recruited for this study were the same subjects participated in the study discussed in Chapter 3. The recruitment adhered to the tenants of the Declaration of Helsinki and approved by the Aston University Research Ethics Committee (2nd of February 2016; #755). All cataract patients were recruited amongst patients visiting the optometry clinic who had significant opacity of their crystalline lens [at least nuclear opalescence (NO) / nuclear colour (NC) grade 3 or posterior subcapsular cataract (P) grade 2 according to Lens Opacities Classification System (LOCS) III (Chylack et al., 1993; Karbassi et al., 1993)] but a healthy fundus. All of them fulfilled the following inclusion criteria:

h) Aged between 20 to 80 years
i) Best corrected visual acuity (BCVA) worse than 6/6 but better or same as 6/18 for the cataractous eye
j) Distance refractive errors were less than ±6DS in sphere, 3.5DC in astigmatism.

The exclusion criteria for this study were as follows:

a) Intraocular pressure (IOP) > 21 mmHg with non-contact tonometer (NCT-510, Nidek, AiChi, Japan)
b) History of intraocular surgery complications
c) Other ocular diseases besides cataract
d) Systemic illness that could affect visual fields (e.g. pituitary lesions, demyelinating diseases, HIV+, AIDS)
k) Uncontrolled diabetes mellitus or untreated hypertension
l) They are pregnant or nursing
m) They were taking any drugs or alcohol that potentially affects reaction time or visual field
n) They have consumed alcohol, nicotine or caffeine less than 2 hours before their perimetric examination

The age-matched normal subjects conformed to the identical inclusion criteria except their BCVA was at least 6/6 and no VF defect in two consecutive VF tests. Both cataract and normal subjects showed no suspicious optic disc appearance (i.e., localized rim loss, optic disc haemorrhage, cup/disc ratio > 0.6 or cup/disc asymmetry > 0.2, notches, localized pallor, or nerve fibre layer defects), and no family history of glaucoma. For subsequent data analyses, all the subjects produced VF results with good reliability indices (FP ≤ 33%, FN ≤ 20%, and FL ≤ 20%) (Anderson and Patella, 1999).

After providing the written informed consent, all normal subjects underwent a comprehensive eye examination to determine their eligibility for this study. On the other hand, cataract subjects were recruited and informed consent was obtained after they were diagnosed in their routine optometric eye examination performed in the optometry clinic.

6.3.2 Methods and Procedures

The procedure was identical to the procedure described in Chapter 3.3.2 (pg. 83). Before proceeding with the VF assessment, all the subjects had undergone comprehensive eye examination including visual acuity test, refraction, non-contact tonometry with Nidek NCT-510 (Nidek, AiChi, Japan), corneal pachymetry and anterior chamber angle imaging using Scheimpflug topographer (TMS-5, Tomey, Nagoya, Japan), external and internal eye examination using slit lamp biomicroscopy and non-mydriatic retinal photography with a fundus camera (Topcon TRC-NW300, Topcon, Tokyo, Japan). The types of cataract were identified by the author using a simple subjective method through slit-lamp biomicroscopy and non-mydriatic fundus camera and then graded according to LOCS III.

VF assessments were only completed after the confirmation of the eligibility of the subjects. All the subjects had undergone at least two visits of the VF tests. Two VF assessments per eye were done in a visit using different threshold strategies i.e. SPARK Precision (SP) from Oculus Twinfield 2 (Oculus, Wetzlar, Germany) and SITA Standard (SS) from Humphrey Field Analyzer (HFA) (Carl Zeiss Meditec, Dublin, CA). There was a ten-minute break interval between VF tests with different strategies. The first visit was used as a familiarization process and an identical order of VF tests was used in both the visits. Central 30° field was tested using both strategies and only reliable VF results in the second visit were used for the data analysis in this study. The details of the VF assessment were described in Chapter 3.3.2.1 and 3.3.2.2. (pg. 84 & 86).

The results of the less severe cataractous eye were selected (better BCVA) if the subject was diagnosed with bilateral cataract. The second tested eye was chosen for the study if there was no difference between both eyes. Only one eye in each patient was used in the data analysis of this study.

A similar analysis method to Chapter 5 was used for the data analysis. The normal subjects recruited were matched with cataract subjects in terms of age and spherical equivalent. Comparison between-strategy was conducted on the MS, MD, and PSD, as well as the testing time in both groups. Correlation and agreement between-strategy of the global indices were also determined. Pointwise analysis of the threshold estimates between-strategy for all the matched 66 test points was also carried out.

6.3.3 Statistical Analysis

Most of the data statistical analysis was performed with IBM SPSS version 22 (IBM Corp, Armonk, NY). Shapiro-Wilk test was used to determine the normality of the data distribution while unpaired t-test or Mann Whitney test was used for the comparison between the groups. Paired t-test or Wilcoxon Signed Rank test was used for the comparison between-strategy. All the correlation tests were used Spearman's correlation coefficient and the agreement between the global indices of the different strategies and also pointwise threshold estimates comparison were conducted using Bland-Altman plots.

6.4 Results

A total of 62 subjects participated in this study which comprised of 31 cataract patients and 31 age-matched healthy control subjects. They had completed at least 2 visits of VF tests. SITA Standard (SS) from HFA and SPARK Precision (SP) from Oculus Twinfield 2 were used in both visits with identical order of tests. The order of the tests was randomized among the subjects. Among the cataract patients, sixteen of them started with SP and 15 started with SS for both of the visits. The interval periods between the two visits were ranged from 1 day to 22 days (median 7 days). Thirteen were males and 18 were females. All cataract subjects were recruited from the patients in the optometry clinic and all confirmations of the diagnosis of cataract were done by the author. All the subjects had also participated in the study described in Chapter 3.

In the current study, only the VF results on their second visit were used for data analysis. All subjects had reliable VF results in their second visit (FP ≤ 33%, FN ≤ 20%, and FL ≤ 20%). The severity levels of the cataract were stratified according to BCVA which are shown in Table 6.1. Among all recruited subjects, 25 of the 31 subjects were diagnosed with nuclear cataract (NO/NC grade 3 and above), three with both nuclear and cortical cataract (C grade 2) and the remaining three had posterior subcapsular cataract (P grade 2) (LOCS III).

Table 6.1: Distribution of cataract severity based on VA

VA	No	%	Type of cataract
6/18	2	6.5	N
6/15	1	3.2	N
6/12	4	12.9	N
6/10	4	12.9	N, PSC
6/9	18	58.1	N, PSC, C
6/7.5	2	6.5	N
Total	31	100.0	

***** N – Nuclear cataract, C – Cortical cataract, PSC – Posterior subcapsular cataract

There were a total of 31 age-matched normal subjects recruited as the control in this study who had also undergone two visits of VF tests using SS and SP with the same testing orders. The interval between visits was ranged from two to 35 days (median seven days). It comprised of 17 males and 14 females. The testing orders were equally divided to 15 started with SP and 16 started with SS. Besides the age, the spherical equivalent of the normal subjects also matched with the cataract subjects (Table 6.2).

Table 6.2: Comparison of age and spherical equivalent between cataract and normal groups

	Cataract				Normal			
	Mean	SD	Median	Range	Mean	SD	Median	Range
Age (yr)	60.8	9.6	63.0	(31–76)	57.5	9.0	60.0	(31–71)
Spherical equivalent (D)	-0.42	1.56	-0.38	(-4.25–2.88)	-1.01	1.71	-0.50	(-5.00–1.50)

Shapiro-Wilk test showed data for the age of cataract patients and SE from the normal subjects was not normally distributed (see Appendix A6.1). Mann-Whitney test was used to ascertain no statistically significant differences of age and SE between cataract and control group (Age: $Z = -1.580$, $p = 0.114$; SE: $Z = -0.493$, $p = 0.622$).

6.4.1 Comparison Within-strategy between Cataract and Age-matched Control Group

The global indices obtained from SS and SP were recorded and tested with the Shapiro-Wilk test to determine the normality of the data distribution (see Appendix A6.2). According to the Shapiro-Wilk test's results, statistical comparisons of global indices between cataract and normal subjects were conducted with the Mann-Whitney test whereas unpaired t-test was used for comparison between-group for testing time spent for the VF test.

Cataract group in this study had shown statistically significant lower values in MS and MD but higher PSD which indicated deeper localised defects as compared to control group (Mann Whitney test: $p = 0.001$ for PSD using SS; $p < 0.001$ for the rest of the indices) (Table 6.3). The normal subjects had also taken significant shorter testing time using SS and SP [Unpaired t-test: (SS) $t = 4.533$, $df = 46.912$, $p < 0.001$; (SP) $t = 5.711$, $df = 60$, $p < 0.001$]. By comparing the SDs of the testing time between the cataract and normal groups, it showed the inter-subject variability in the data of testing time with SP was impressively consistent as compared to the fluctuation shown with SS.

Table 6.3: Comparison of visual field data between cataract and age-matched control group

		Cataract				Normal				Unpaired t test*/Mann Whitney test
		Mean	SD	Median	Range	Mean	SD	Median	Range	p
MS (dB)	SS	26.17	3.65	27.18	(13.45–30.29)	30.03	1.17	30.27	(27.45–32.55)	<0.001
	SP	27.08	3.05	27.97	(17.30–30.97)	30.36	1.00	30.05	(28.06–32.02)	<0.001
MD (dB)	SS	-2.72	2.62	-1.38	(-10.95–0.72)	0.53	1.05	0.77	(-2.21–1.79)	<0.001
	SP	0.31	2.50	1.52	(-6.55–2.94)	2.71	1.23	3.05	(-2.22–3.96)	<0.001
PSD (dB)	SS	2.52	1.49	2.07	(1.14–8.19)	1.68	0.29	1.66	(1.17–2.54)	0.001
	SP	1.60	0.61	1.51	(0.81–3.37)	1.13	0.21	1.12	(0.79–1.72)	<0.001
Testing time (min)	SS	7.63	1.36	7.37	(5.72–11.12)	6.37	0.75	6.15	(5.05–8.18)	<0.001*
	SP	3.67	0.13	3.65	(3.42–3.92)	3.47	0.15	3.45	(3.23–3.82)	<0.001*

6.4.2 Comparison Between-strategy within Cataract and Age-matched Control Group

The global indices and testing time of cataract and age-matched control group were compared between the strategies (Table 6.4). According to the normality of the data distribution, Wilcoxon Signed Rank test was used for the comparison between-strategy of MS, MD, and PSD in cataract group and also MD in control group whereas paired t-test was used for comparison between-strategy of MS and PSD in control group and also testing time in both groups. Mean sensitivity (MS) of SP was significantly higher than MS of SS in the cataract group but not in the control group ($p > 0.05$). However, the difference of mean MS between-strategy in cataract subjects was less than 1 dB which was clinically insignificant. SPARK Precision (SP) had consistently produced much elevated MD in either cataract patients or normal subjects which was translated as less diffuse loss detected by SP. Significant lower PSD was also shown with SP compared to using SS in either cataract patients or controls subjects but a larger difference

was observed in cataract patients. There was significant shorter testing time with SP which saved 52% and 46% of the testing time compared to using SS in cataract and normal group respectively.

Table 6.4: Comparison between SITA Standard and SPARK Precision

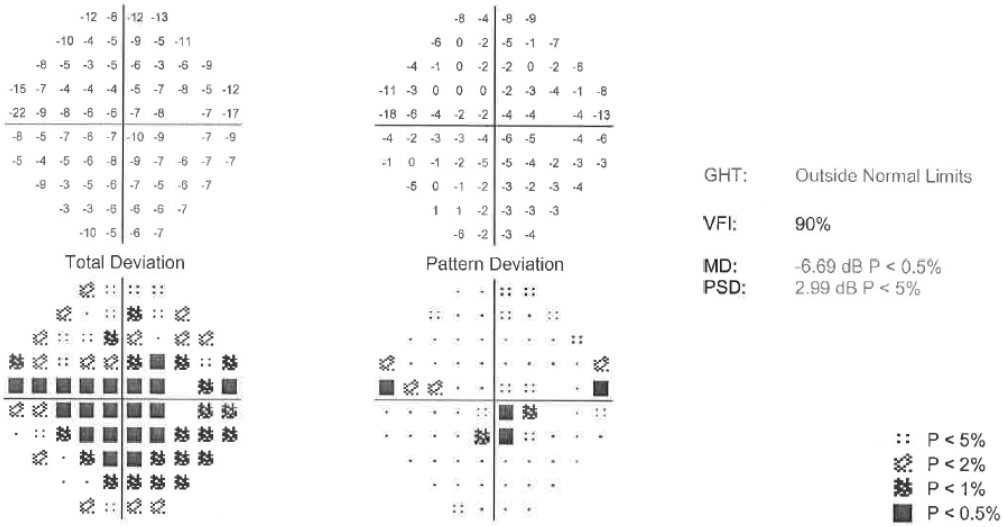
		SITA Standard (SS)		SPARK Precision (SP)		Paired t-test/ Wilcoxon Signed Rank test		
		Mean^/ Median	SD^/ (Range)	Mean^/ Median	SD^/ (Range)	Z/t^	df	p
MS (dB)	C	27.18	(13.45-30.29)	27.97	(17.3-30.97)	-2.962		0.003
	N	30.03^	1.17^	30.36^	1.00^	-1.860^	30	0.073*
MD (dB)	C	-1.38	(-10.95-0.72)	1.52	(-6.55-2.94)	-4.860		<0.001
	N	0.77	(-2.21-1.79)	3.05	(-2.22-3.96)	-4.782		<0.001
PSD (dB)	C	2.07	(1.14-8.19)	1.51	(0.81-3.37)	-4.507		<0.001
	N	1.68^	0.29^	1.13^	0.21^	10.088^	30	<0.001*
Test time (min)	C	7.64^	1.36^	3.67^	0.13^	16.940^	30	<0.001*
	N	6.37^	0.75^	3.47^	0.15^	23.609^	30	<0.001*

C - Cataract; N - Normal
*Paired t test

Two examples of the results from SP and SS were used to show differences between the strategies. The first example shown in Figure 6.1 demonstrates both strategies detected the diffuse loss in a cataractous eye which was graded as NC/NO5 with LOCS III and BCVA was 6/18. As shown in the results, SP displays shallower field defect with a lesser number of abnormal total deviations points showing $p < 0.5\%$ and lower MD as compared to SS. The example in Figure 6.2 showed that the VF results of another cataractous eye which was graded as NC/NO4 and C2 with LOCS III and BCVA was 6/9. Its diffuse loss of the sensitivity was not exhibited distinctly through the result of SP as compared to the result of SS. There are the same numbers of the abnormal total deviation (TD) and pattern deviation (PD) points showed by the result of SP. The depressed normal hill of vision in SP could mask the severity of the diffuse loss which had shown lower MD than SS's MD. Another 5 cases were found showing

similar findings as shown in Figure 6.2 which SP was not able to detect the diffuse loss found with SS.

SITA Standard



* MS of the matching 66 test points = 22.09 dB

SPARK Precision

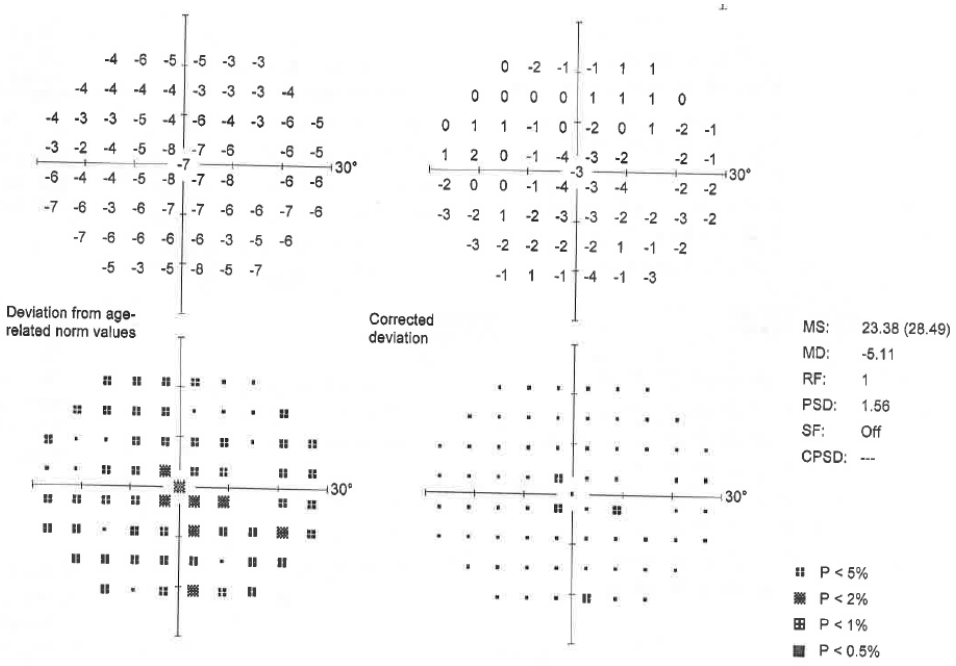
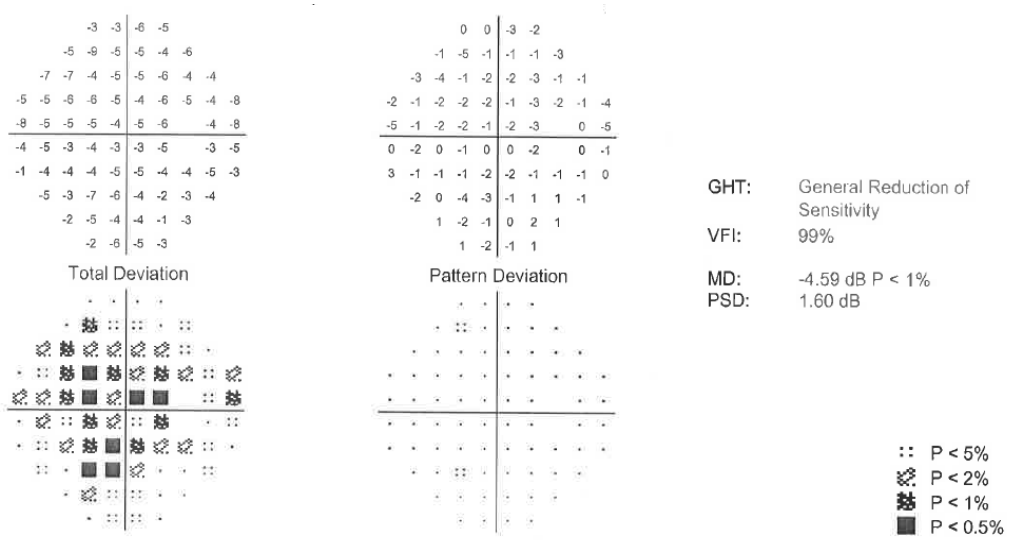


Figure 6.1: Results from SS (upper) and SP (lower) that show apparently diffuse loss detected in the VF of an eye with nuclear cataract (NC/NO5 using LOCS III) and BCVA of

6/18. The number of the abnormal points in total deviation probability map is markedly higher than the number in pattern deviation probability map for both strategies.

SITA Standard



* MS of the matching 66 test points = 23.70 dB

SPARK Precision

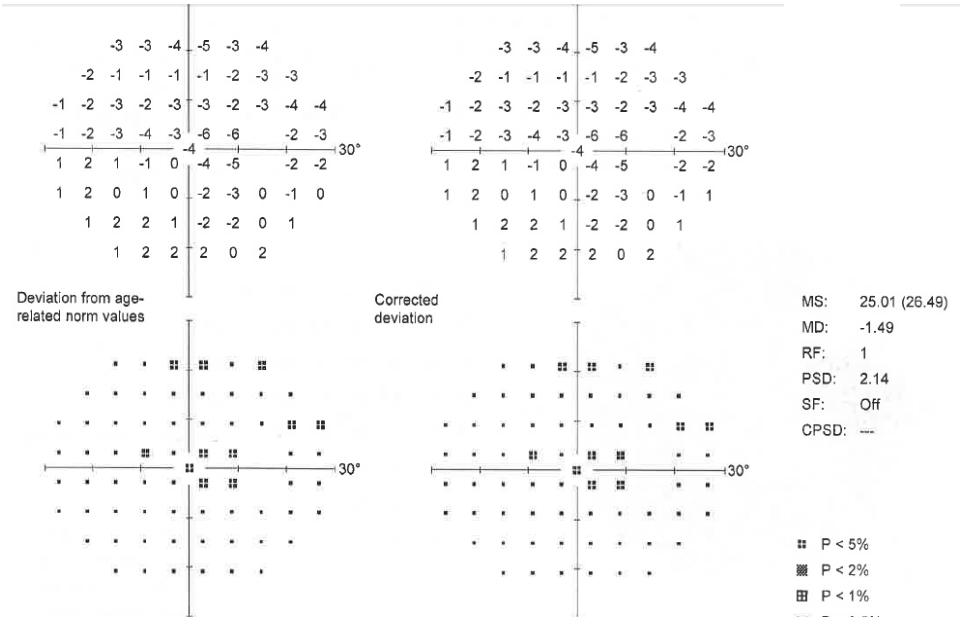


Figure 6.2: Diffuse loss in VF is not shown prominently with SP (lower) compared to the result from SS (upper) in an eye with nuclear and cortical cataract (NC/NO4 and C2 using LOCS III) and BCVA of 6/9. The number of the abnormal points shown in total

deviation probability map equals to the number of abnormal points in pattern deviation probability map of SP. It is distinctly more diffuse loss shown with the result of SS

6.4.3 Agreement and Correlation Between-strategy

6.4.3.1 Mean sensitivity

The data from all subjects (cataract and age-matched control subjects) was used to determine the correlation of MS between strategies and a statistically significant correlation (Spearman correlation coefficient: rho = 0.855, $p < 0.001$) was found between the strategies. Statistical agreement between the MS of both strategies was determined with the Bland-Altman plot as shown in Figure 6.3. More than 95% data points (59/62 = 95.2%) were located within LoA. The bias/mean difference (95% LoA) was 0.61 dB (-2.30, 3.53 dB) in all subjects but the bias could increase with lower MS as indicated by the statistically significant proportional bias (t=-3.006, p=0.004) with regression test.

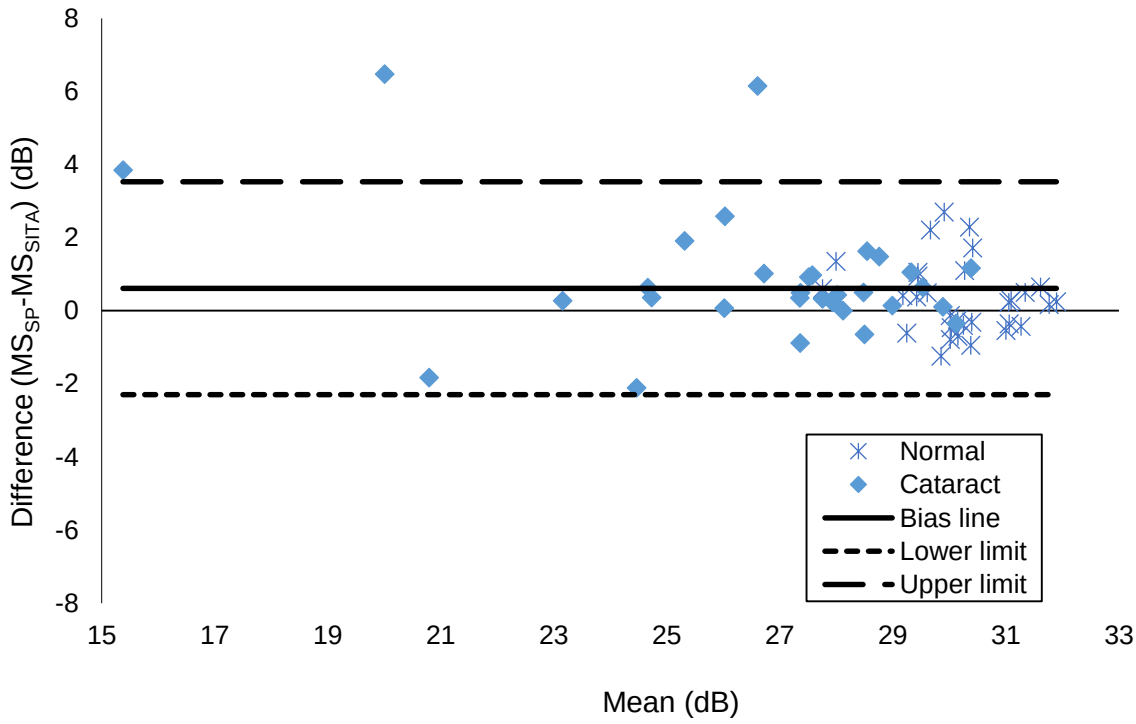


Figure 6.3: Bland-Altman plot of mean sensitivity between SS and SP in all subjects

The agreement between MS of the strategies was also evaluated using Bland-Altman plots in cataract and control group respectively. In the cataract group, it showed that 93.5% of data points (29/31) were located within 95% LoA with no proportional bias was found (t=-1.937, p=0.062) (Figure 6.4). The bias (95% LoA) had increased to 0.90 dB (-2.69, 4.50 dB) in this cataract group. Whereas in control group, similar result with 93.5% of data points (29/31) was found located within the 95% LoA without proportional bias (t = -1.1079, p = 0.289) (Figure 6.5). The bias (95% LoA) had become smaller with 0.33 dB (-1.60, 2.25 dB).

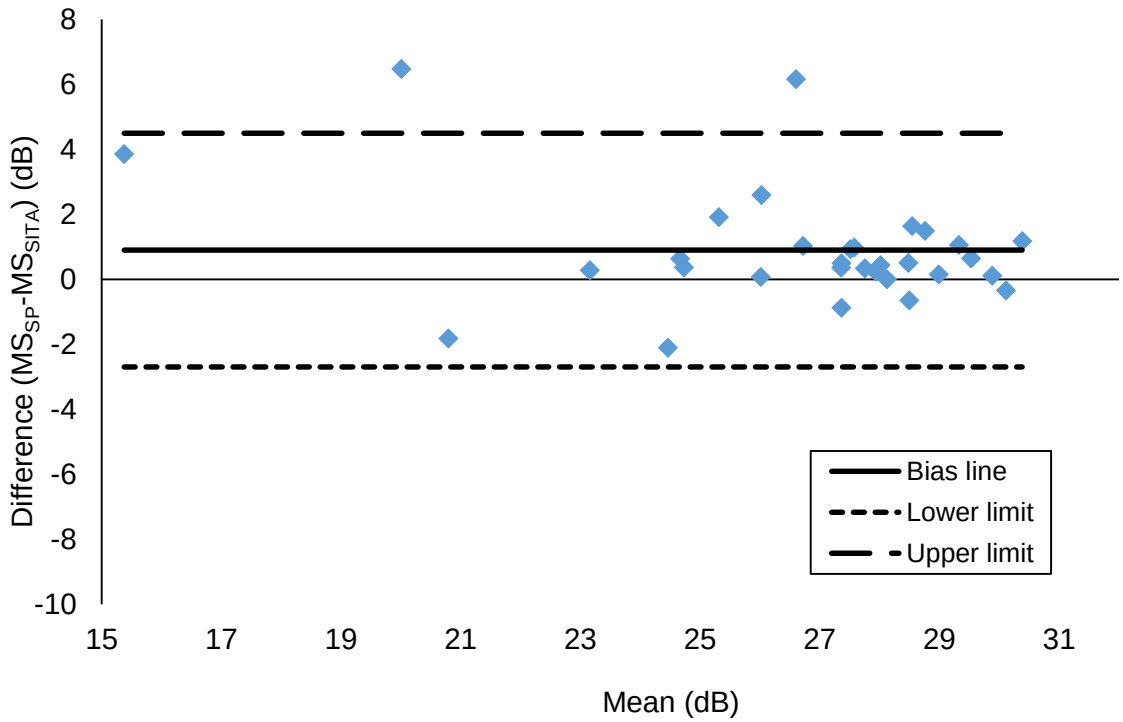


Figure 6.4: Bland-Altman plot of mean sensitivity (MS) between using SS and SP in cataract group

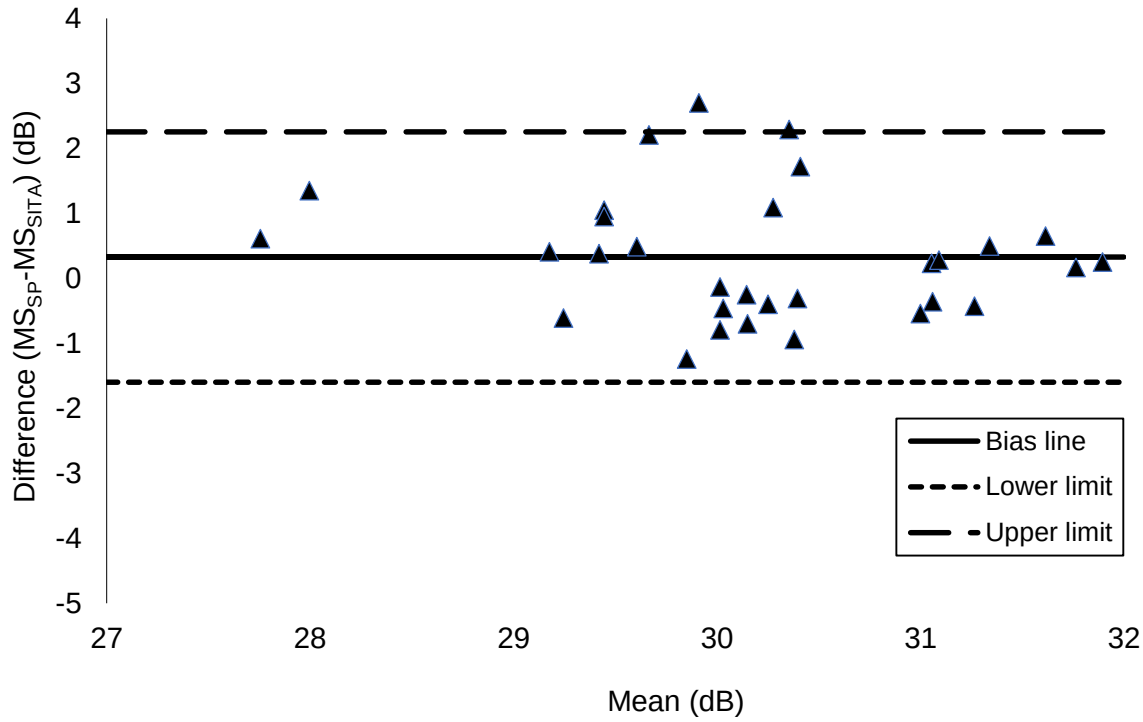


Figure 6.5: Bland-Altman plot of mean sensitivity between using SS and SP in age-matched control group

6.4.3.2 Mean deviation

Mean deviations (MD) from both strategies were also found to have a statistically significant correlation (Spearman correlation coefficient: rho = 0.881, $p < 0.001$) in all subjects. Bland-Altman plots of MD between SS and SP were conducted separately for all subjects, cataract patients and age-matched control subjects respectively (Figure 6.6, 6.7 and 6.8). The percentages of data points located within LoA achieved more than 95% in the plots of all subjects (95.2%) and control subjects only (96.8%). Whereas in the group of cataract patients, it was found only 93.5% of the data points were located within the LoA. The bias (95% LoA) found was 2.60 dB (-0.01, 5.22 dB) in all subjects and 2.18 dB (0.38 dB, 3.98 dB) in age-matched control group. If only referring to cataract subjects only, the bias (95% LoA) was higher at 3.03 dB (-0.01 dB, 6.06 dB).

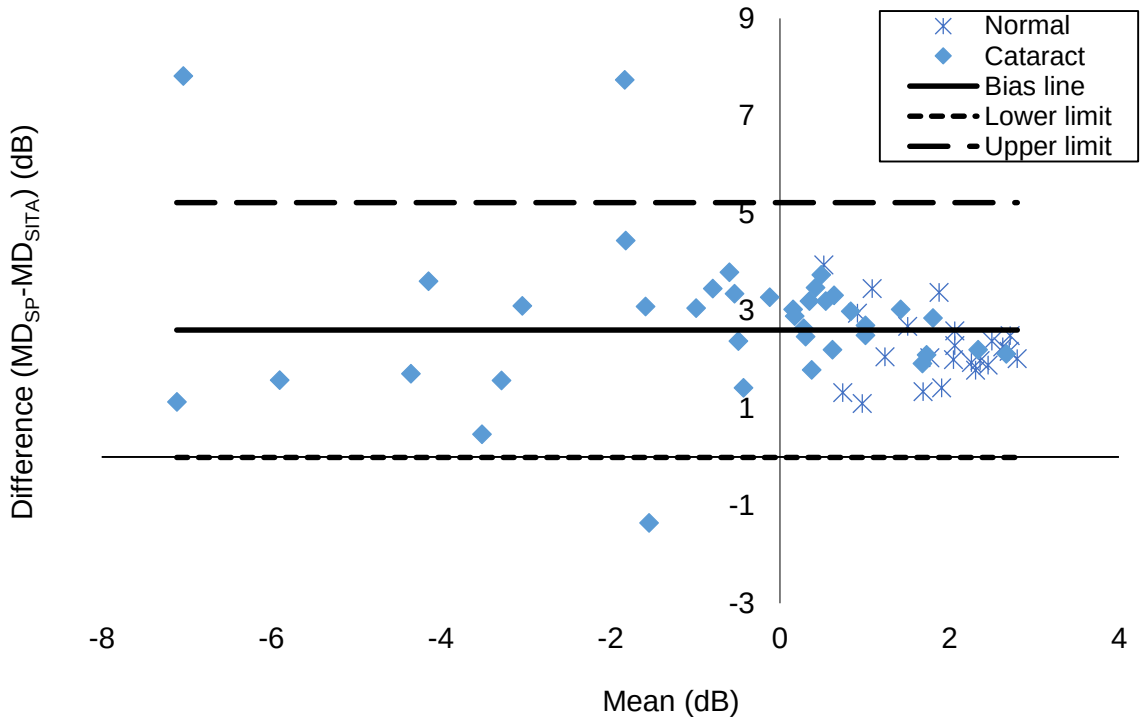


Figure 6.6: Bland-Altman plot of mean deviation (MD) between SS and SP in all subjects (cataract patients and age-matched normal subjects)

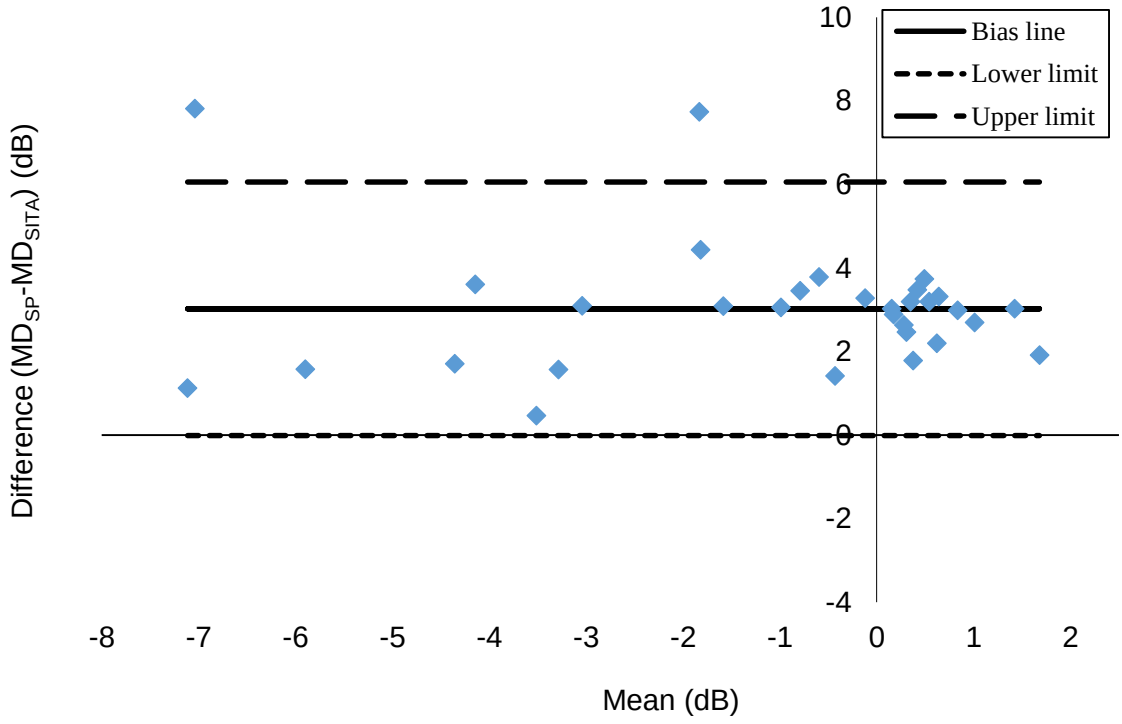


Figure 6.7: Bland-Altman plot of mean deviation between SS and SP in cataract patients

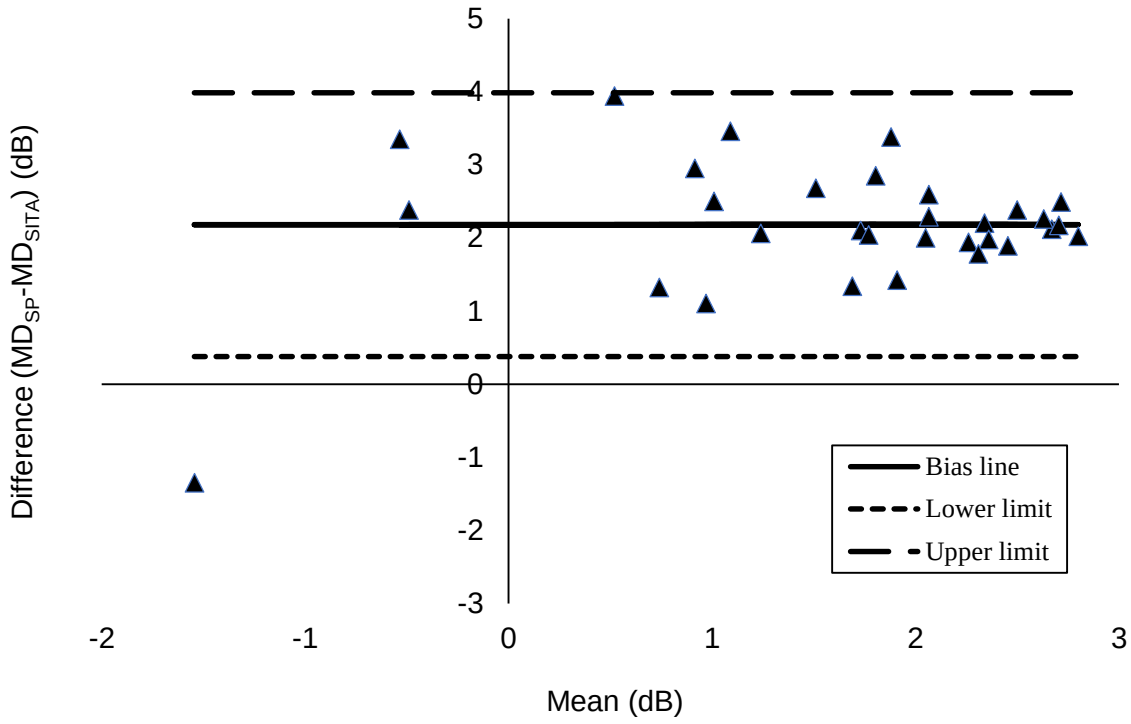


Figure 6.8: Bland-Altman plot of mean deviation between SS and SP in age-matched control subjects

6.4.3.3 Pattern standard deviation

Pattern standard deviations (PSD) from both strategies were also found to be positively correlated in the group comprised of cataract patients and age-matched control subjects (Spearman correlation coefficient: rho = 0.543, $p < 0.001$). Bland-Altman plot showed 96.8% of the data points (60/62) were located within 95% LoA in the combined group of cataract patients and normal subjects (Figure 6.9). The bias (95% LoA) between the strategies according to the data recruited was -0.73 dB (-2.38 dB, 0.91 dB) but it could increase with higher PSD according to the significant proportional bias shown in the data ($t = -10.562$, $p < 0.001$).

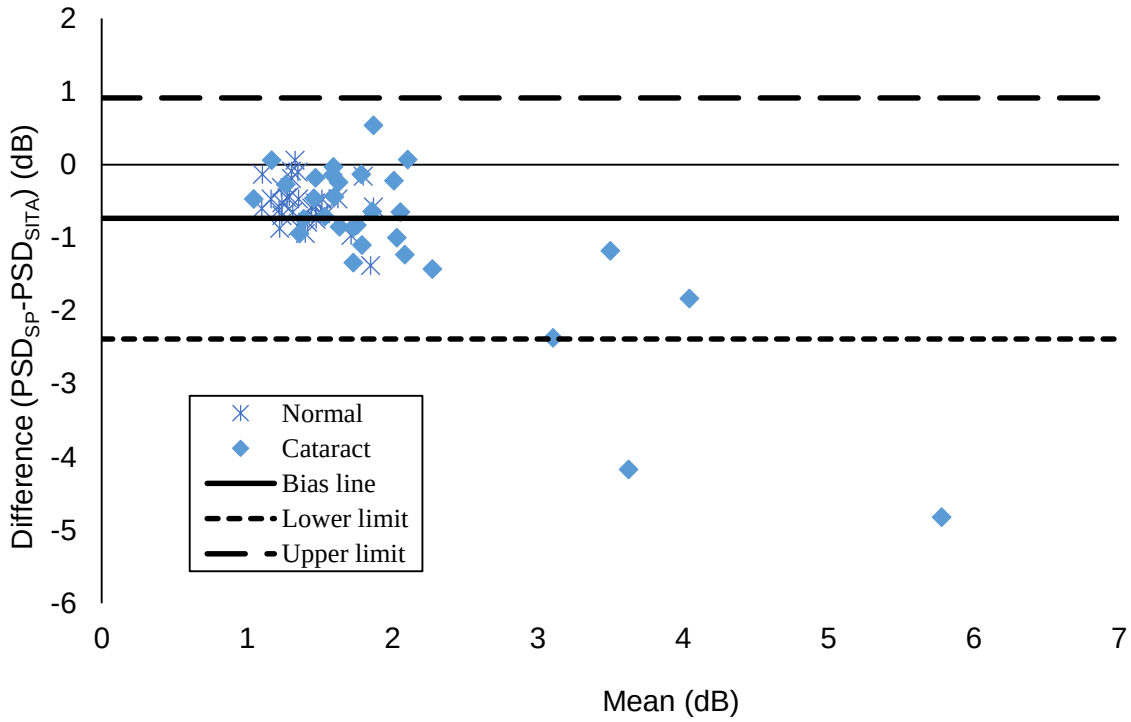


Figure 6.9: Bland-Altman plot of pattern standard deviation between SS and SP in cataract and age-matched control group

The agreement of PSD between SS and SP was also evaluated in the cataract group according to Bland-Altman plot with 6.5% of data points that were found located outside 95% LoA (Figure 6.10) and the bias also increases with higher PSD ($t = -8.101$, $p < 0.001$). The bias (95% LoA) of PSD between strategies was -0.92 dB (-3.14 dB, 1.29 dB) in this cataract group.

.

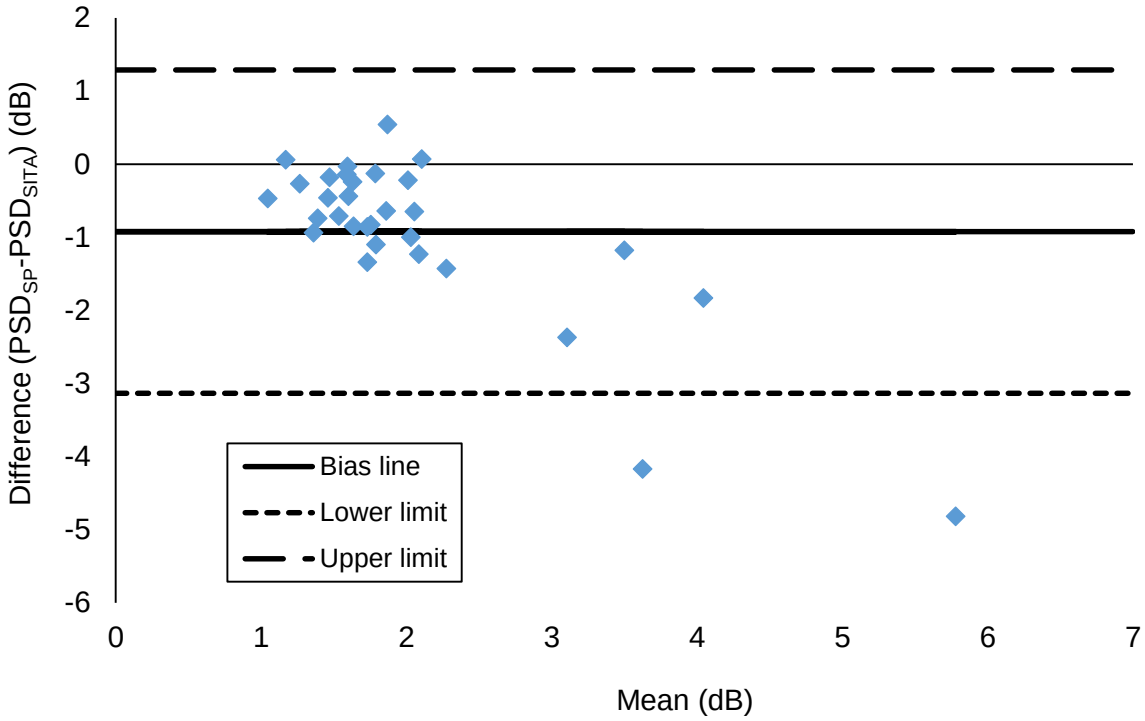


Figure 6.10: Bland-Altman plot of pattern standard deviation between SS and SP in cataract patients

In age-matched control group, only 93.5% of data points were located within the 95% LoA (Figure 6.11) even though the bias (95% LoA) in PSD has reduced to -0.55 dB (-1.14 dB, 0.04 dB). No significant proportional bias was observed in Bland-Altman plot (t = -1.915, p = 0.065).

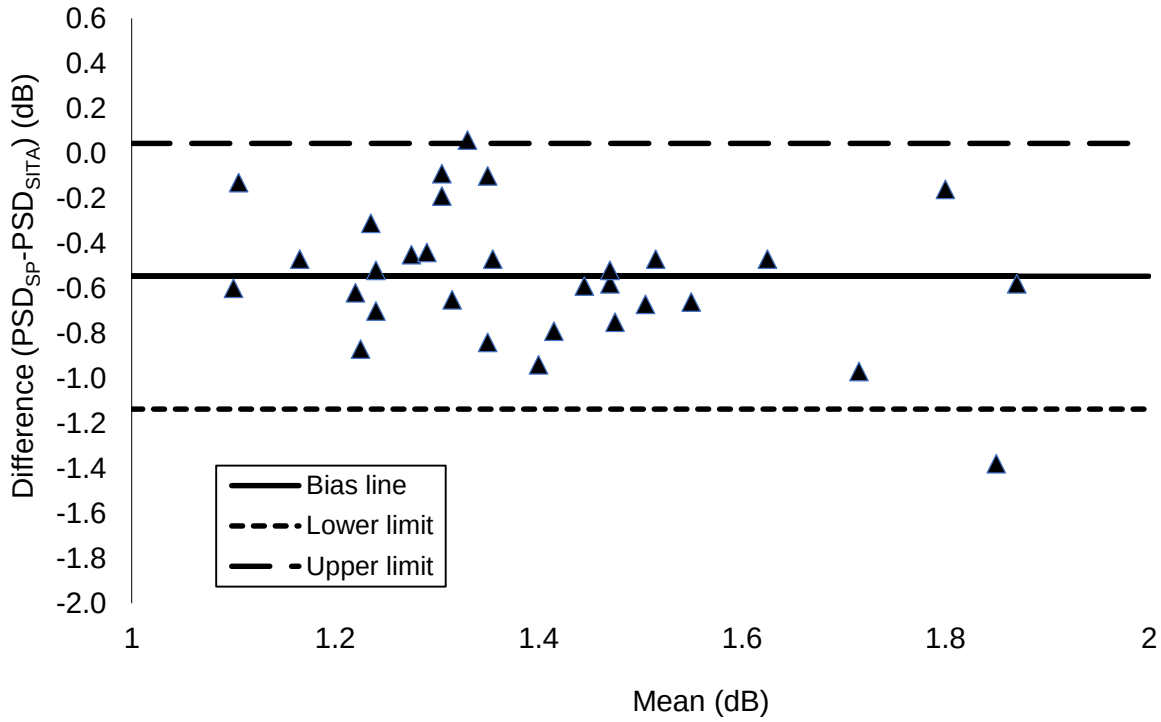


Figure 6.11: Bland-Altman plot of pattern standard deviation between SS and SP in age-matched control subjects (vs Cataract)

6.4.4 Pointwise analysis of comparison between-strategy in cataract patients

The pointwise analysis was conducted using Bland-Altman plots at each of the 66 test locations where the biases and 95% LoAs of the threshold estimates were determined (Figure 6.12). Interestingly, the biases between the strategies were higher (threshold estimate of SP was higher than SS) mainly at the edge points of superior-nasal area of the central field whereas lower biases (threshold estimate of SS higher than SP) were found more on a cluster of test points located at the other side of the field i.e. inferior-temporal area of the central field. It showed an almost similar pattern for the bias changes found in glaucoma subjects in the study discussed in Chapter 5 (pg. 154). The values of the bias were progressively reduced from nasal to temporal.

		2.84 8.84	**3.23** 10.29	**1.74** 7.76	**2.42** 7.72	**1.45** 6.13	**1.90** 7.23		
	2.77 9.34	**1.39** 8.96	**0.87** 5.56	**1.74** 4.41	**0.87** 3.03	**1.03** 5.60	**0.26** 4.41	**0.65** 6.13	
2.13 9.05	**1.61** 5.79	**1.29** 6.28	**0.55** 5.74	**0.58** 3.64	**0.71** 3.59	**0.52** 3.64	**0.26** 4.44	**0.10** 4.68	**-0.06** 5.21
3.42 12.37	**2.26** 7.14	**0.90** 5.69	**0.13** 3.94	**0.19** 4.68	**0.13** 4.32	**1.29** 3.96		**-0.29** 3.55	**-0.71** 6.41
2.32 7.99	**1.77** 8.89	**0.77** 4.29	**-0.23** 3.02	**0.48** 2.99	**0.65** 3.60	**0.48** 4.26		**-0.29** 3.89	**-1.23** 4.63
2.23 9.49	**1.13** 7.17	**1.26** 7.92	**0.23** 4.55	**0.84** 3.40	**0.52** 4.17	**0.19** 4.57	**-0.48** 3.64	**-0.58** 3.71	**-0.48** 5.72
	1.77 10.07	**0.32** 6.00	**1.06** 6.88	**0.68** 4.22	**0.06** 4.23	**-0.52** 6.62	**-1.29** 3.86	**-1.06** 4.98	
		1.71 7.96	**1.45** 8.00	**0.65** 6.04	**-0.29** 4.91	**-0.16** 5.50	**0.26** 4.41		

Bias (dB)
<-1.00
-1.00 ~ -0.01
0.00~0.99
1.00~1.99
≥2.00

Figure 6.12: Bland-Altman agreement analysis between SS and SP at the 66 test points in cataract subjects (as the field view of right eye). Bias (upper and bolded number) and the 95% limits of agreement (lower and unbolded number) in the threshold estimates for each 66 matching test points between SS and SP.

*The bias is the MS of SP minus that of SS
*The colours represent the ranges of the bias in MS between SS and SP

6.5 Discussion

The current study was aimed to compare the capability of the threshold strategy, SPARK Precision (SP) of Oculus Twinfield 2 to detect a reduction of the sensitivity in VF caused by cataract with SITA Standard (SS) of HFA. All otherwise healthy cataract patients' results were compared to the age-matched normal subjects. In view of reduction of threshold sensitivity with aging (Hermann et al., 2008; Heijl et al., 1987; Jaffe et al., 1986; Zulauf et al., 1994) and higher refractive error (Aung et al., 2001; Rudnicka and Edgar, 1996), the normal subjects

recruited as controls were matched for the group mean age and spherical equivalent to the cataract group.

The cataract subjects had shown higher average MS with SP than with SS and lack of the agreement of MS between the strategies was found in cataract subjects. Besides the lack of data points located within 95% LoA, the LoA between the strategies of 7.19 dB had indicated a bad agreement between SS and SP in the cataract subjects according to the criteria used by Luithardt et al. (2015). The bias observed also showed relatively larger than the bias in glaucoma patients (0.90 dB vs 0.37 dB) as shown in the result from Chapter 5 (pg. 146). This is probably due to both strategies (SITA and SPARK) work better in sensitivity estimation for glaucomatous eyes on the ground that both utilized previous samples of glaucoma results to set up the algorithms (Bengtsson et al., 1997; Gonzalez de la Rosa and Gonzalez-Hernandez, 2013).

Overestimation of MD was again produced by SP with an even larger margin when compared to MD from SS in the cataract subjects. Elevated normal hill of vision in SP field analysis masked the diffuse field loss caused by the cataract. The agreement of MD was not established between the two strategies in cataract subjects. The bias of MD was more than 3 dB or LoA > 5 dB and it is considered as “bad agreement” according to the criteria from Luithardt et al. (2015). The agreement was also not achieved in the normal group which was also found to have similar bias and LoA of MD in Chapter 5. The bias more than 2 dB for MD was unacceptable for the agreement between-strategy in normal subjects even though more than 95% of data points located within 95% LoA in Bland-Altman plot. A marked increase of the MD difference between-strategy relative to the difference in MS was also observed in the cataract group and It was similarly found in normal and glaucoma subjects (Chapter 5) which was discussed in Chapter 5 (pg. 161). Weighting method and age-matched norm values could be playing a part for the drastic increase of the differences.

Despite reduced PSD was found in cataract subjects as compared to glaucoma patients in the previous chapter (see Chapter 5, pg. 144), SP still significantly underestimated PSD as compared to using SS despite showing much smaller irregularities in SP. The bias of PSD between SS and SP increased with a deeper localised defect in the cataract group. Similarly, the increment pattern was also shown in the glaucoma group (Chapter 5, pg. 152). The agreement of PSD between the two strategies in the cataract group was considered to be acceptable according to the bias (-0.92 dB) and LoA (4.43 dB) but it lacked 95% of data points located within the 95% LoA. With the bias increases with deeper localised field defect, the agreement will not be achieved if more eyes with severe media opacity were recruited. However, a good agreement of PSD between the strategies was achieved in normal group disregard the lack of 95% data points within LoA which could be improved with a larger sample size for normal subjects.

By comparing the results of SS and SP, there was no statistically significant difference of MS between the strategies in the normal subjects. The influence of reduced fatigue effect was surprisingly not shown here even though SP saved 46% of the testing time compared to SS. It was contradicting with the statistical result shown in Chapter 4 (pg. 119). This is most probably due to much larger sample size was recruited hence higher statistical power was obtained in the previous chapter. Moreover, a broader age-range of the normal group and younger subjects were recruited in the previous study (Chapter 4). Good agreement of MS between the strategies shown in the normal subjects with a bias of only 0.33 dB and LoA of 3.85 dB according to clinical evaluation criteria provided by Luithardt et al. (2015). Even though only 93.5% of the data points fell within the LoA, it should be able to increase if a larger number of healthy normal subjects were recruited.

The cataract group exhibited a profound reduction of MS as well as MD compared to the age-matched normal subjects with using either SS or SP. The results of the present study are in agreement with previous research which has also shown reduction of MD due to the loss of

lens transparency (Rehman Siddiqui et al., 2007; Koucheki et al., 2004; Smith et al., 1997; Hayashi et al., 2001; Kook et al., 2004; Musch et al., 2006; Ang et al., 2010; Kim et al., 2001) whereas PSD was also found to be significantly higher in the cataract group than normal control group. The opacities in the cataractous eye could be sufficiently deep to cause significant field defect in pattern deviation map (Phu et al., 2017) which inevitably increase the PSD of the visual field. However, the difference between the cataract and control group apparently was smaller when using SP. Pattern standard deviation (PSD) is known to be unaffected by cataract (Lam et al., 1991; Vijaya et al., 2005; Ang et al., 2010; Rehman Siddiqui et al., 2007). The average PSD in the cataract group (1.6 dB) was still lower than 1.8 dB which is recommended as the diagnostic cut-off point for abnormal field defect with SP (Gonzalez de la Rosa et al., 2013). The testing time used for cataract subjects was unsurprisingly longer compared to normal subjects when using SS as SITA strategy duration is associated with the severity of field defect (Wild et al., 1999a; Roggen et al., 2001). SPARK Precision also required longer testing time in cataract subjects than in normal subjects ($p < 0.001$) even though it was more consistent (smaller SD) and the difference of test duration between groups was comparatively much smaller when using SP. In Chapter 5 (pg. 143), SP had also produced slightly longer testing time in glaucoma patients but consistent time was observed even though the range of the glaucoma severity on patients recruited was wide.

Pointwise analysis of the comparison between-strategy for MS in each 66 test locations has shown that higher bias test points (overestimated threshold sensitivity of SP) were located more on the nasal side of the field and lower bias test points (underestimated threshold sensitivity of SP) were located on the temporal side of the field. It debunks the assumption made in Chapter 5 that SP tends to overestimate threshold sensitivity on the common areas of glaucomatous field defect. The trend of bias decreasing from nasal to temporal was found in both cataract and glaucoma groups. It was also found in normal subjects that displayed a larger average bias of MS in the nasal field compared to the temporal field (Chapter 4, pg. 124). The test points of the nasal field which are located further from the optic disc/blind spot

as compared to the points in temporal field could be associated with the trend of the bias reduction (Rao et al., 2017). Large LoA, especially for test points in the nasal field, further indicates that threshold estimates from both strategies were definitely not interchangeable. Further investigation for this finding is required.

In brief, according to the evaluation of the agreement between SS and SP, it was found that in the presence of media opacity, the agreement was not achieved for all global indices i.e. MS, MD, and PSD even though there was lack of severe cataractous eyes with BCVA worse than 6/9 in this study. Large bias was observed especially for MD even in the normal subjects. SPARK Precision (SP) is not recommended to be a clinical alternative to SS for the subjects with media opacities or exhibit general loss of sensitivity.

Nevertheless, SP saved around 52% of the testing time in cataract subjects which was also approximately the same amount of time-saving found in glaucoma patients (Chapter 5, pg. 144). It could be used as a screening tool but not to replace SITA strategies.

As the prevalence of cataract increases with age (Na et al., 2014; Nowak and Smigielski, 2015; McCarty et al., 1999; Klein et al., 1992a), it was a struggle to find older subjects with healthy eyes that can match the age of the cataract group. More than 60% of the cataract subjects recruited had BCVA of at least 6/9. In spite of continuous efforts made to find more subjects with lower BCVA, the accessibility and affordability of the cataract surgery in the urban area of Petaling Jaya (the research location) had hindered the recruitment of eyes with more severe cataract. Most of the recruited cataract patients in this study were under monitoring without the need to proceed for cataract surgery.

As visual acuity is not sufficient to evaluate the severity of the cataract (Shandiz et al., 2011; Pesudovs and Elliott, 2003; Elliott et al., 1989), the use of LOCS III (Chylack et al., 1993; Karbassi et al., 1993) for the subjective clinical grading of cataracts helped to determine more

precisely the severity level of the cataract. Nevertheless, the intention of this study was mainly to compare the performance of SP and SS in threshold estimation, thus it did not proceed to determine the correlation of the cataract severity level to VF result. Moreover, pure forms of the cataract are rarely found and commonly the heterogeneity of the cataract could also affect the subjective grading by experienced clinicians (Lam et al., 1991) or instruments (Yao and Flammer, 1993).

6.6 Conclusion

The cataract patients exhibited a significant reduction of sensitivity with either SS or SP. The MS for the matching test points and MD were reduced with both SS and SP but SP showed a smaller margin of the difference between the cataract and normal groups.

The MS of SP was higher than SS with an average 0.90 dB of bias which may be benefited from more than 50% saving of the testing time but a poor agreement between the strategies was found. It showed large LoA and a lack of 95% of data points located within 95% LoA. The bad agreement was also displayed for the MD between the two strategies of which the bias between the strategies has markedly increased and masked the diffuse loss of sensitivity induced by the cataract.

The PSD in the cataract subjects was higher than the control group either using SS or SP even though the index was less affected by cataract relative to glaucoma. However, PSD was underestimated with SP and the bias of PSD as compared to SS increased with the deeper localized defect. The agreement of PSD between the strategies was not accepted due to the fact that lack of data points was found within the 95% LoA despite relatively lower bias and LoA were found compared to MD's.

SPARK Precision (SP) was found more likely to overestimate the sensitivities in the nasal field and underestimate the points in the temporal field with regards to SS's result. Similar patterns

were found in the glaucoma group and normal control group as well. It is unlikely related to the common glaucomatous areas but indeed a systemic difference may exist between the strategies. Before the difference is further rectified, SP is not recommended to be used for eyes with either diffuse loss or localized field loss as a clinical replacement to SS.

CHAPTER 7

GENERAL SUMMARY OF THE RESULTS, CONCLUSIONS, LIMITATIONS OF THE STUDY AND PROPOSALS FOR FUTURE WORKS

7.1 General Summary of the Results and Conclusions

The introduction of SPARK was aimed to reduce the fluctuation in SAP by shortening the testing time. A narrower confidence interval for normality could be obtained and hence enables VF defect to be detected earlier statistically. The use of data from more than 90,000 VF results enables the statistical relationships between the different testing points in central field to be utilised. With the established relationships, a shorter testing time is achieved. A brevity VF examination could be beneficial to reduce patient's fatigue, inter-subject and test-retest variability. It may also minimize the influence of the learning effect and improve the stability of the result. The series of these studies were intended to evaluate the performance of SPARK in achieving the aforementioned advantages by between-visit analysis and between-strategy comparison to the "gold standard", SITA.

The stability of the results from SPARK was evidenced by the between-visit comparative study within the same strategy in glaucoma patients and normal subjects. However, the same consistency of SPARK was not presented in the cataract patients in this study possibly due to the high proportion of perimetric-inexperienced cataract patients recruited. Prior experience in perimetry had helped to improve the sensitivity of the VF either using SITA or SPARK. The impact of the perimetric experience was shown to be mitigated in SPARK which has the advantages of shorter testing time and extra training phase through the first phase of the strategy test. The stability of SPARK was also shown by the absence of the improvement of MS with SPARK between the visits within normal subjects either with or without perimetric experience whereas it was otherwise with SITA. The evidence of stability was further shown with lower inter-test variability in SPARK which most likely benefited from the use of average

threshold estimates and the rejection of the most extreme estimate in each test point by SPARK algorithm. However, the lack of sensitivity improvement between-visit in the peripheral field was found with SPARK as compared to the central field which could be partly due to the sensitivity reduction of a cluster of test points in the superior field possibly related to the lower-positioning of the surveillance camera. Further rectification is required to identify this shortcoming of SPARK.

A reliable VF result with SPARK was easier to be achieved as compared to SITA in either normal subjects, subjects with focal loss (i.e. glaucoma) or diffuse loss (i.e. cataract). SPARK presented with a higher number of reliable results which was mainly attributed to the monitoring methods of fixation loss rate and shorter test duration. Repeating VF test and delaying of the diagnostic decision could be avoided with SPARK especially novices in perimetric test. Prior experience in perimetry was not found to be associated with the reliability of the results with SPARK but it was not the case with SITA in glaucoma patients.

In brief, the first series of the studies showed that SPARK displayed a lower inter-test variability, a lower number of unreliable results and was less influenced by the perimetric-experience of subjects, all of which are desirable improvements compared to SITA. Nevertheless, it also showed a familiarisation process of the VF test was inevitably required for the following series of the studies.

About 40% of time-saving was found with SPARK which will be the liking of the patients but higher average threshold sensitivity of close to 1 dB was estimated as compared to SITA in the between-strategy comparative study in normal subjects. It could have resulted from the fading of the fatigue effect. However, the differences in the instrumentation and algorithms applied could also be part of the reason for the higher threshold estimates of SPARK. Most of the test points in SPARK had a higher threshold estimate than SITA especially in the superior region and the pointwise bias between the two strategies showed a progression of reduction

from superonasal towards inferotemporal of the VF which yet to be rectified. Nevertheless, SPARK still produced the average threshold estimates that considerably achieved an acceptable agreement with SITA's estimates. Furthermore, the differences between threshold estimates of the two strategies were found to be decreased towards old age in healthy subjects.

The performance of SPARK was further evaluated in patients diagnosed with glaucoma which is characterised by VF focal loss. Its results were evidently not comparable to SITA despite more testing time being saved and lower average bias of estimated sensitivity was found. However, an agreement could not be reached in global indices between the two strategies which was exhibited by the large LoA using Bland-Altman analysis. The high bias of more than 2 dB and large LoA were found in MD between the strategies and a noticeable reduction of PSD with SPARK was also observed in glaucoma patients and even in healthy subjects. The overestimation of total deviation which could be due to the depressed normal hill of vision in normative data indeed produced lower AGIS scores. Shallower VF defect was found with SPARK even though the defect size was not different with SITA which underestimated the severity level of VF defect. As a result, poor diagnostic sensitivity was found especially for early glaucoma at least based on these studies results. In spite of a wider range of threshold sensitivity with more depressed sensitivities found in glaucoma patients, the pointwise biases between SPARK and SITA displayed the changes from the higher biases in the nasal field to the lower biases in the temporal field. The variation pattern of bias was found unexplainably similar to the one found in normal subjects.

The performance of SPARK in patients with diffuse loss (i.e. cataract) was below expectation as presented in the final series of these studies. Even though the cataract patients presented more depressed general sensitivity with much reduced focal loss (lower PSD) as exhibited in glaucoma patients, agreements between SPARK and SITA were not achieved in either MS, MD or PSD. Higher overestimation of MD with SPARK was continued to be shown with even

higher bias than the one reported in glaucoma patients from the previous study. It was again evidenced that the diffuse loss by cataract could be missed by SPARK which possibly masked by the depressed normal hill of vision in the normative data. The pointwise biases of the threshold estimates between SPARK and SITA were again showing the similar pattern of change with nasal field displayed the highest positive bias while temporal field had the most negative value.

In conclusion, SPARK algorithm was able to show an acceptable agreement between threshold sensitivities with SITA but it was restricted to normal subjects. Furthermore, the threshold estimates consistently showed the same variations of bias between strategies in the studies involving subjects with different VF defects possibly demonstrated a systemic difference between the strategies. A questionable and possibly depressed normative database used in the SPARK algorithm affects the validity of the global indices and consequently worsen the diagnostic sensitivity of SPARK strategy. SPARK was yet to be established as a clinical alternative to SITA despite a marked shorter test duration and reduced inter-test variability.

7.2 Limitation of the study

In order to achieve a sufficient number of subjects for the study within the period permissible for a doctorate degree, the recruitment of the glaucoma patients was conducted through a few ophthalmology clinics which were reliant on the subjective assessment of the different ophthalmologists. The diagnosis of early glaucoma could be varied among the ophthalmologists especially when there was no significant VF defect found. A more standardized and objective diagnostic method such as structural diagnosis using OCT (Dong et al., 2016; Lisboa et al., 2012) or Heidelberg retinal tomograph (HRT) (Heidelberg Engineering, Heidelberg, Germany) (Maslin et al., 2015) will be more ideal for the confirmation of glaucoma. However, it was not possible in this study considering the aforementioned instruments were not available in the participated centres.

Recruitment of the patients with early glaucoma is preferred for the determination of the diagnostic sensitivity between different algorithms but the diagnostic accuracy between the two methods depends on the VF defects severity distribution among the patients (Petraco et al., 2018) and each level should be recruited to have a better statistical analysis. This study was having a dilemma of whether to recruit patients with early stage of glaucoma or patients with different severity levels of glaucoma. Ultimately a sample size with more than 70% with early stage of glaucoma was recruited and the rest were patients with more severe glaucoma in this study. Even though glaucoma is characterized by focal VF loss and cataract is diffuse VF loss, there were also moderate to advanced stage of glaucoma patients displayed significant diffuse loss of VF. Some cataract subjects who may not have a generalised media opacity but a certain part of the cataract could be dense enough to result in depressed PSD. Subjects with purely diffuse VF loss or focal VF loss were difficult to sort after which could be helped by only recruiting the subjects with early stage of the disease.

The learning effect may not be totally nullified even though there was an additional visit for the familiarization process provided for all the patients. It was shown in Chapter 3 that higher sensitivity was found in patients with perimetric experienced even in the second visit. Extra sessions of the perimetric examination are necessary to have a better comparison between the strategies as stated in Chapter 4 as well (pg. 130).

SITA Fast (SF) perhaps is more appropriate to compare with SPARK due to both have similar test duration (Abreu-Gonzalez et al., 2018) but SITA Standard (SS) was chosen ahead of SF in this study mainly due to SF was not recommended for monitoring patients with confirmed VF loss (Artes et al., 2002) whereas SS is the more common standard strategy used for a diagnosed glaucoma patient with better repeatability (Sekhar et al., 2000). Moreover, the VF tests conducted on the glaucoma patients participated in this study were intended to be used for the monitoring of their glaucoma progression. The next following study could proceed with

the comparison between SPARK and SF or the newly introduced fast threshold strategy, SPARK Faster.

7.3 Future works

7.3.1 The normative database of SPARK

In Chapter 5 and 6, SPARK was shown to produce distinctly shorter testing durations which resulted in vaguely higher MS but with a clinically subtle bias (< 1 dB) as compared to SS. However, remarkably larger differences were found in Total Deviation and Pattern Deviation between the two strategies which potentially were associated with age-matched normative database applied in SPARK despite differences in MD calculation and numbers of stimulus points were also partly at fault. Furthermore, the average MD of normal subjects was markedly higher than the value reported by Gonzalez de la Rosa et al. (2013). Hence, the normative database used by SPARK needs to be reviewed especially with Twinfield perimeter considering the use of simulated maximum stimulus illumination for the estimation of sensitivity. It was not clear whether Easyfield or Twinfield was used for the compilation of the database and how it was compiled. The selection of the “normal” subjects, the number of the examinations used and how to minimise fatigue and learning effect are not standardised across all the VF algorithms. It was arbitrarily determined by each manufacturer and not obliged to publish. Nevertheless, it requires enormous resources and time for the development of the database which needs to be considered by the manufacturer. Ideally, each instrument should have an individual database especially in the case of Twinfield and Easyfield which have a difference in stimulus luminance range. A large sample size of healthy subjects comprising of stratified age groups from at least 20 to 80 years old is required to set up the normative database for a strategy test.

Besides the doubt in regards to the normative database, the systemic difference between SS and SP needs to be further identified which was exhibited as the asymmetrical pointwise difference of MS between the strategies in this study. It could start with the comparison

between SPARK and SITA within a narrower age range of normal subjects. Different age groups should be recruited which possibly helps to determine whether any contribution of age to the asymmetrical pointwise differences between the strategies. The correlation between the pointwise bias and the distances from the blind spot/optic disc could be further determined to verify the possible association as mentioned in Chapter 6 (pg. 195). Perhaps, the systemic difference could also present by the comparison between SPARK and SITA Fast or even SITA Faster if it is due to the methodological difference between the algorithms.

7.3.2 The learning effect of SPARK

The SPARK Precision consists of four phases including the first phase which takes only 40 s which could also be used as a training phase. The training phase can be run alone by using SPARK Training. Its effectiveness in minimizing the impact of the learning effect should be further evaluated in groups that exhibited distinctly different VF defects. Only subjects that are naive to perimetric examination should be recruited and multiple visits up to at least 5 visits should be completed for all subjects in order to identify the number of tests required to have stable threshold estimates.

The reduction of the threshold estimates of the superior field in the following visit which was shown in Chapter 3 should be re-investigated. The effect of the lower-than-eye-level camera used in Oculus Twinfield could be assessed with another strategy to rule out its effect on the variability between-visit found in Chapter 3. The camera is not installed in the central aspect of the field to allow the fixation loss to be monitored through the central threshold. The benefit of such an installation may need to be reassessed. The learning effect of SPARK can be further determined whether it is transferable to other perimetry which requires longer testing time such as SITA. If that is possible, the use of the SPARK Training is preferred rather than a full VF test for the familiarization process. The SPARK Training can also be investigated whether it can be used as a screening tool for various VF loss as it has a much shorter testing duration than a suprathreshold test that usually used for screening. Subjects with different VF defects

can be recruited to perform the VF test with SPARK Training in order to determine its accuracy and sensitivity with regards to the types or locations of the VF defects.

7.3.3 The diagnostic sensitivity and detection of the progression of visual field defect

If the normal hill of vision applied in the total deviation has been reviewed and improved, the diagnostic sensitivity of SPARK of different types of glaucoma including ocular hypertension patients should be conducted. A short VF test which also provides a stable threshold estimation is always a key goal for tools in visual function assessment. It allows more frequent perimetric examinations which indeed improves the sensitivity of an analytic program to detect VF defect progression. Stable VF results can also minimize the possibility of VF defect progression masked by learning effect. The early stage of glaucoma could be progressed diffusely (Gonzalez de la Rosa et al., 2010). Thus the detection of the diffuse loss is important to improve the sensitivity in diagnosing progression. SPARK has been emphasized on producing stable MD. A progression analytic program recommended by the manufacturer, the threshold less-noise trend (TNT) program (Gonzalez de la Rosa et al., 2010a) analyses the changes of MD. The use of SPARK with TNT should also be evaluated to determine their efficacy to measure glaucoma progression.

LIST OF REFERENCES

ABREU-GONZALEZ, R., GONZALEZ-HERNANDEZ, M., PENA-BETANCOR, C., PALOMA RODRIGUEZ-ESTEVE, P., GONZALEZ DE LA ROSA, M. (2018). New visual field indices of disharmony for early diagnosis of glaucoma, alone or associated with conventional parameters. *European Journal of Ophthalmology*, 28(5), p.590-597

AKAR, Y., YILMAZ, A., YUCEL, I. (2008). Assessment of an effective visual field testing strategy for a normal pediatric population. *Ophthalmologica*, 222, p.329-333.

ALLINGHAM, R.R., DAMJI, K.F., FREEDMAN, S.F., MOROI, S.E., RHEE, D.J., SHIELDS, M.B. (2011). *Shields Textbook of Glaucoma*. 6th ed. Philadelphia: Lippincott Williams & Wilkins.

ANCTIL, J.L. and ANDERSON, D.R. (1984). Early foveal involvement and generalized depression of the visual field in glaucoma. *Archives of Ophthalmology*, 102(3), p.363-370.

ANDERSON, A.J. (2003). Spatial resolution of the tendency-oriented perimetry algorithm. *Investigative Ophthalmology and Visual Science*, 44, p.1962-1968.

ANDERSON, A.J. and JOHNSON, C.A. (2006). Comparison of the ASA, MOBS and ZEST threshold methods. *Vision Research*, 46, p.2403-2411.

ANDERSON, D.R., PATELLA, V.M. (1999). *Automated Static Perimetry.* 2nded.St. Louis, MO: Mosby.

ANDERSON, R.S., REDMOND, T., McDOWELL, D.R., BRESLIN, K.M., ZLATKOVA, M.B. (2009). The robustness of various forms of perimetry to different levels of induced intraocular stray light. *Investigative Ophthalmology and Visual Science*, 50, p.4022-4028.

ANG, G.S., SHUNMUGAM, M., AZUARA-BLANCO, A. (2010). Effect of cataract extraction on the Glaucoma Progression Index (GPI) in glaucoma patients. *Journal of Glaucoma*, 19(4), p. 275-278.

AOKI, Y., TAKAHASHI, G., KITAHARA, K. (2007). Comparison of Swedish interactive threshold algorithm and full threshold algorithm for glaucomatous visual field loss. *European Journal of Ophthalmology*, 17(2), p.196-196.

ARAIE, M., YAMAGAMI, J., SUZIKI, Y. (1993). Visual field defects in normal-tension and high-tension glaucoma. *Ophthalmology*, 100, p.1808-1814.

ARMALY, M.E. (1971). Visual field defects in early open angle glaucoma. *Transactions of the American Ophthalmological Society*, 69, p.147-162.

ARMALY, M.F. (1969). The size and location of the normal blind spot. *Archives of Ophthalmology*, 81(2), p.192-201.

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

ARTES, P.H., IWASE, A., OHNO, Y., KITAZAWA, Y., CHAUHAN, B.C. (2002). Properties of perimetric threshold estimates from Full Threshold, SITA Standard, and SITA Fast strategies. *Investigative Ophthalmology and Visual Science*, 43, p.2654-2659.

ARTES, P.H., NICOLELA, M.T., LEBLANC, R.P., CHAUHAN, B.C. (2005). Visual field progression in glaucoma: total versus pattern deviation analyses. *Investigative Ophthalmology and Visual Science*, 46, p.4600-4606

ASBELL, P.A., DUALAN, I., MINDEL, J., BROCKS, D., AHMAD, M., EPSTEIN, S. (2005). Age-related cataract. *Lancet*, 365, p.599-609

ASMAN, P. and HEIJL, A. (1992). Evaluation of methods for automated Hemifield analysis in perimetry. *Archives of Ophthalmology*, 110, p.820-826.

ASMAN, P. and HEIJL, A. (1992a). Glaucoma Hemifield Test. Automated visual field evaluation. *Archives of Ophthalmology*, 110, p.812-819.

ASMAN, P. and HEIJL, A. (1992b). Weighting according to location in computer-assisted glaucoma visual field analysis. *Acta Ophthalmologica*, 70, p.671-678.

ASMAN, P. and HEIJL, A. (1994). Diffuse visual field loss and glaucoma. *Acta Ophthalmologica (Copenh)*, 72, p.303-308

ATALAY, E., NONGPIUR, M.E., YAP, S.C., WONG, T.T., GOH, D., HUSAIN, R., PERERA, S.A., AUNG, T. (2016). Pattern of visual field loss in primary angle-closure glaucoma across different severity levels. *Ophthalmology,* 123(9), p.1957-1964.

ATCHISON, D.A. (1987). Effect of defocus on visual field measurement. Ophthalmic *and Physiological Optics*, 7, p.259-265.

AULHORN, E. and HARMS, H. (1972). Visual perimetry. IN: JAMESON, D., HURVICCH, L.M.(eds.) *Handbook of Sensory Physiology*: Volume VII/4: Visual Psychophysics. Berlin, Germany: Springer-Verlag, p.102-145.

AULHORN, E. and KARMEYER, H. (1977). Frequency distribution in early glaucomatous visual field defects. *Documenta Ophthalmologica Proceedings Series*, 14, p.75-83.

AUNG, T., FOSTER, P.J., SEAH, S.K., Chan, S.P., Lim, W.K., Wu, H.M., Lim, A.T., Lee, L.L., Chew, S.J. (2001). Automated static perimetry: the influence of myopia and its method of correction*. Ophthalmology*, 108(2), p.290-295.

AUTZEN, T. and WORK, K. (1990). The effect of learning and age on short-term fluctuation and mean sensitivity of automated state perimetry. *Acta Ophthalmologica*, 68, p.327-330.

AYDIN, A., KOCAK, I, AYKAN, U., CAN, G., SABAHYILDIZI, M., ERSANLI, D. (2015). The influence of the learning effect on automated perimetry in a Turkish population. *Journal français d'ophtalmologie*, 38(7), p.628-632.

BAL, T., COECKELBERGH, T., VAN LOOVEREN, J., ROZEMA, J.J., TASSIGNON, M.J. (2011). Influence of cataract morphology on straylight and contrast sensitivity and its relevance to fitness to drive. *Ophthalmologica*, 225, p.105-111.

BALLON, B.J., ECHELMAN, D.A., SHIELDS, M.B., OLLIE, A.R. (1992). Peripheral visual field testing in glaucoma by automated kinetic perimetry with the Humphrey Field Analyzer. *Archives of Ophthalmology*, 110(12), p.1730-1732.

BARALDI, P., ENOCH, J.M., RAPHAEL, S. (1987). A comparison of visual impairment caused by nuclear (NC) and posterior subcapsular (PSC) cataracts. IN: GREVE, E.L. and HEIJL, A. (eds.) Proceedings of the Seventh International Visual Field Symposium,

Amterdam, September 1986. Documenta Ophthalmologica Proceedings Series. The Hague, The Netherlands: Martius Nijhoff/Dr W Junk Publishers; Documenta Ophthalmologica Proceedings Series, 49, p.363-366.

BARDE, M.P. and BARDE, P.J. (2012). What to use to express the variability of data: Standard deviation or standard error of mean? *Perspectives in Clinical Research*, 3(3), p.113-116.

BARKANA, Y., GERBER, Y., MORA, R., LIEBMANN, J.M., RITCH, R. (2006). Effect of eye testing order on automated perimetry results using the Swedish Interactive Threshold Algorithm Standard 24-2. *Archives of Ophthalmology*, 124(6), p.781-784.

BARLOW, H.B. (1958). Temporal and spatial summation in human vision at different background intensities. *The Journal of Physiology*, 141(2), p.337-350.

BEBIE, H., FANKHAUSER, F., SPAHR, J. (1976).Static perimetry: strategies. *Acta Ophthalmologica (Copenh)*, 54(3), p.325-338.

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(5), p.519-522.

BENGTSSON, B. and HEIJL, A. (1998). Evaluation of a new perimetric threshold strategy, SITA, in patients with manifest and suspect glaucoma. *Acta Ophthalmologica Scandinavica*, 76(3), p.268-272.

BENGTSSON, B. and HEIJL, A. (1998a). SITA Fast, a new rapid perimetric threshold test. Description of methods and evaluation in patients with manifest and suspect glaucoma. *Acta Ophthalmologica Scandinavica*, 76(4), p.431-437.

BENGTSSON, B. and HEIJL, A. (1999). Comparing significance and magnitude of glaucomatous visual field defects using SITA and Full Threshold strategies. *Acta Ophthalmologica Scandinavica*, 77(2), p.143-146.

BENGTSSON, B. and HEIJL, A. (1999a). Inter-subject variability and normal limits of the SITA standard, SITA Fast, and the Humphrey Full Threshold computerized perimetry strategies, SITA STATPAC. *Acta Ophthalmologica Scandinavica*, 77(2), p.125-129.

BENGTSSON, B. and HEIJL, A. (2000). False-negative responses in glaucoma perimetry: indicators of patient performance or test reliability? *Investigative Ophthalmology and Visual Science* 41(8), p.2201-2204.

BENGTSSON, B. and HEIJL, A. (2006). Diagnostic sensitivity of fast blue-yellow and standard automated perimetry in early glaucoma: a comparison between different test programs. *Ophthalmology*, 113(7), p.1092-1097.

BENGTSSON, B., HEIJL, A. and OLSSON, J. (1998). Evaluation of a new threshold visual field strategy, SITA, in normal subjects*. Acta Ophthalmologica Scandinavica*, 76(2), p.165-169.

BENGTSSON, B., LINDGREN, A., HEIJL, A., LINDGREN, G., ASMAN, P., PATELLA, M. (1997a). Perimetric probability maps to separate change caused by glaucoma from that caused by cataract. *Acta Ophthalmologica Scandinavica*, 75(2), p.184-188.

BENGTSSON, B., OLSSON, J., HEIJL, A., ROOTZÉN, H. (1997). A new generation of algorithms for computerized threshold perimetry, SITA. *Acta Ophthalmologica Scandinavica*, 75(4), p.368-375.

BENITEZ-DEL-CASTILLO, J., REGI, T., MOTA, I., GARCIA-SANCHEZ, J. (2008).Correlation between structure and function in glaucoma. *Investigative Ophthalmology & Visual Science*, 49(13), 731. (ARVO Annual Meeting Abstract)

BERGIN, C., REDMOND, T., NATHWANI, N., VERDON-ROE, G.M., CRABB, D.P., ANDERSON, R.S., GARWAY-HEATH, D.F. (2011). The effect of induced intraocular straylight on perimetric tests. *Investigative Ophthalmology and Visual Science*, 52(6), p.3676-3682.

BERNARDI, L., COSTA, V.P., SHIROMA, L.O. (2007). Flicker perimetry in healthy subjects: influence of age and gender, learning effect and short-term fluctuation. *Arquivos brasileiros de oftalmologia*, 70(1), p.91-99.

BETTELHEIM, F.A. and CHYLACK, L.T. (1985). Light scattering of whole excised human cataractous lenses. Relationships between different light scattering parameters. *Experimental Eye Research*, 41, p.19-30.

BICKLER-BLUTH, M., TRICK, G.L., KOLKER, A.E., COOPER, D.G.(1989). Assessing the utility of reliability indices for automated visual fields. Testing ocular hypertensives. *Ophthalmology*, 96(5), p.616-619.

BIRT, C.M., SHIN, D.H., SAMUDRALA, V., HUGHES, B.A., KIM, C., LEE, D. (1997). Analysis of reliability indices from Humphrey visual field tests in an urban glaucoma population. *Ophthalmology*, 104(7), p.1126-1130.

BIZIOS, D., HEIJL, A., BENGTSSON, B. (2011). Integration and fusion of standard automated perimetry and optical coherence tomography data for improved automated glaucoma diagnostics. *BMC Ophthalmology*, 11, p.20.

BLAND, J.M. and ALTMAN, D.G. (1999). Measuring agreement in method comparison studies. *Statistical Methods in Medical Research*, 8(2), p.135-160.

BLOCH, A. (1885). Expèriencessur la vision. Paris. Société de Biologie Mémoires, 37, p.493-495.

BLUMENTHAL, E.Z., SAMPLE, P.A., ZANGWILL, L., LEE, A.C., KONO, Y., WEINREB, R.N. (2000). Comparison of long-term variability for standard and short-wavelength automated perimetry in stable glaucoma patients. *American Journal of Ophthalmology*, 129(3), p.309-313.

BLUMENTHAL, E.Z., SAMPLE, P.A., BERRY, C.C., LEE, A.C., GIRKIN, C.A., ZANGWILL, L., CAPRIOLI, J., WEINREB, R.N. (2003). Evaluating the several sources of variability for standard and SWAP visual fields in glaucoma patients, suspects, and normals. *Ophthalmology*, 110(10), p.1895-1902.

BOEGLIN, R.J., CAPRIOLI, J., ZULAUF, M. (1992). Long-term fluctuation of the visual field in glaucoma. *American Journal of Ophthalmology*, 113(4), p.396-400.

BOLAND, M.V., ZHANG, L., BROMAN, A.T**.,** JAMPEL, H.D., QUIGLEY, H.A. (2008). Comparison of optic nerve head topography and visual field in eyes with open-angle and angle-closure glaucoma. *Ophthalmology*, 115(2), p.239-245.e2

BOURNE, R.R., TAYLOR, H.R., FLAXMAN, S.R., KEEFFE, J**.**, LEASHER, J**.**, NAIDOO, K**.**, PESUDOVS, K**.**, WHITE, R.A**.**, WONG, T.Y**.**, RESNIKOFF, S**.**, JONAS, J.B**.**; Vision Loss Expert Group of the Global Burden of Disease Study. (2016). Number of people blind or visually impaired by glaucoma worldwide and in world regions: a meta-analysis. *PLoS One,* 11, p.e0162229.

BOURNE, R.R**.**, STEVENS, G.A**.**, WHITE, R.A**.**, SMITH, J.L**.**, FLAXMAN, S.R**.**, PRICE, H**.**, JONAS, J.B**.**, KEEFFE, J**.**, LEASHER, J**.**, NAIDOO, K**.**, PESUDOVS, K**.**, RESNIKOFF, S**.**,TAYLOR, H.R**.**; Vision Loss Expert Group. (2013). Causes of vision loss worldwide, 1990-2010: a systematic analysis. *The Lancet. Global health.* 1(6), p.e339**-**349.

BOZKURT, B., YILMAZ, P.T., IRKEC, M. (2008). Relationship between Humphrey 30-2 SITA Standard test, Matrix 30-2 threshold test, and Heidelberg retina tomograph in ocular hypertensive and glaucoma patients. *Journal of Glaucoma*, 17(3), p.203-210.

BRAIS, P. and DRANCE, S.M. (1972). The temporal field in chronic simple glaucoma. *Archives of Ophthalmology*, 88(5), p.518-522.

BRENTON, R.S. and PHELPS, C.D. (1986). The normal visual field on the Humphrey Field Analyzer. *Ophthalmologica*, 193, p.56-74.

BRENTON, R.S., PHELPS, C.D., ROJAS, P., WOOLSON, R.F. (1986). Interocular differences of the visual field in normal subjects. *Investigative Ophthalmology and Visual Science*, 27, p.799-805.

BROMAN, A.T., QUIGLEY, H.A., WEST, S.K., KATZ, J., MUNOZ, B., BANDEEN-ROCHE, K., TIELSCH, J.M., FRIEDMAN, D.S., CROWSTON, J., TAYLOR, H.R., VARMA, R., LESKE, M.C., BENGTSSON, B., HEIJL, A., HE, M., FOSTER, P.J.(2008). Estimating the rate of progressive visual field damage inthose with open-angle glaucoma, from cross-sectional data*. Investigative Ophthalmology and Visual Science*, 49(1), p.66-76.

BRUDER, G. and KIETZMAN, M. (1973). Visual temporal integration for threshold, signal detectability, and reaction time measures. *Attention, Perception, and Psychophysics*, 13(2), p. 293-300.

BRUSINI, P. and FILACORDA, S. (2006). Enhanced Glaucoma Staging System (GSS 2) for classifying functional damage in glaucoma. *Journal of Glaucoma,* 15(1), p.40-46.

BRUSINI, P. and JOHNSON, C.A. (2007). Staging functional damage in glaucoma: review of different classification methods. *Survey of Ophthalmology*, 52, p.156-179.

BUDENZ, D.L., FEUER, W.J., ANDERSON, D.R. (1993). The effect of simulated cataract on the glaucomatous visual field. *Ophthalmology,* 100(4), p.511-517.

BUDENZ, D.L., RHEE, P., FEUER, W.J., McSOLEY, J., JOHNSON, C.A., ANDERSON, D.R. (2002). Sensitivity and specificity of the Swedish interactive threshold algorithm for glaucomatous visual field defects. *Ophthalmology*, 109(6), p.1052-1058.

BUDENZ, D.L., RHEE, P., FEUER, W.J., McSOLEY, J., JOHNSON, C.A., ANDERSON, D.R. (2002a). Comparison of glaucomatous visual field defects using standard full threshold and Swedish interactive threshold algorithms. *Archive of Ophthalmology*, 120(9), p.1136-1141

BURGANSKY-ELIASH, Z., WOLLSTEIN, G., PATEL, A., BILONICK, R.A., ISHIKAWA, H., KAGEMANN, L., DILWORTH,W. D., SCHUMAN, J.S. (2007). Glaucoma detection with matrix and standard achromatic perimetry. *British Journal of Ophthalmology*, 91(7), p.933-938

CALVO PEREZ, P., GIL-ARRIBAS, L., FERRERAS, A., OTIN, S., ALTEMIR, I., FERNANDEZ, S., PABLO JULVEZ, L., FUERTES, I. (2010). Relationship between flicker FDF perimetry and standard automated perimetry. *Acta Ophthalmologica*, 88, s246.

CAMP, A.S. AND WEINREB, R.N. (2017). Will perimetry be performed to monitor glaucoma in 2025? *Ophthalmology*, 124, p.S71-S75.

CAPRIOLI, J. and SPAETH, G.L. (1984). Comparison of visual field defects in the low-tension glaucomas with those in the high-tension glaucomas. *American Journal of Ophthalmology*, 97, p.730-737.

CAPRIOLI, J. and SPAETH, G.L. (1985). Static threshold examination of the peripheral nasal visual field in glaucoma. *Archives of Ophthalmology*, 103(8), p.1150-1154.

CAPRIOLI, J., SEARS, M., MILLER, J.M. (1987). Patterns of early visual field loss in open-angle glaucoma. *American Journal of Ophthalmology*, 103(4), p.512-517.

CAPRIS, P., AUTUORI, S., CAPRIS, E., PAPADIA, M. (2008). Evaluation of threshold estimation and learning effect of two perimetric strategies, SITA Fast and CLIP, in damaged visual fields. *European Journal of Ophthalmology*, 18(2), p.182-190.

CAPRIS, P., GATTI, G., CORALLO, G., ROMITI, S., GANDOLFO, E. and ZINGIRIAN, M. (1999). Comparing SITA and Full Threshold strategies. Perimetry Update 1998/1999, pp. 11–15 Proceedings of the XIIIth International Perimetric Society Meeting, Gardone Riviera (BS), Italy, September 6–9, 1998. Hague: Kugler Publications.

CARL ZEISS MEDITEC. (2010). *Humphrey Field Analyzer II-i Series*, *User manual*: Dublin, CA: Carl Zeiss Meditec.

CARRILLO, M.M., ARTES, P.H., NICOLELA, M.T., LEBLANC, R.P., CHAUHAN, B.C. (2005). Effect of cataract extraction on the visual fields of patients with glaucoma. *Archives of Ophthalmology*, 123, p.929-932.

CASSON, R.J. and JAMES, B. (2006). Effect of cataract on frequency doubling perimetry in the screening mode. *Journal of Glaucoma*, 15(1), p.23-25.

CASTRO, D.P., KAWASE, J., MELO, L.A.Jr. (2008). Learning effect of standard automated perimetry in healthy individuals. *Arquivos Brasileiros de Oftalmolohgia*, 71(4), p. 523-528.

CHAGLASIAN, M. (2013). Sharpen your visual field interpretation skills. *Review of Optometry*, Feb 15, 2013. Available from: **http://www.reviewofoptometry.com/content/d/glaucoma/c/39563/** [Assessed 4Oct, 2015]

CHAKRAVARTI, T. (2017). Assessing precision of Hodapp-Parrish-Anderson criteria for staging early glaucomatous damage in an ocular hypertension cohort: a retrospective study. *Asia Pacific Journal of Ophthalmology*, 6(1), p.21-27

CHAPLIN, G.B.B., EDWARDS, J.H., GEDYE, J.L., MARLOWS, S. (1973). Automated system for testing visual fields. *Proceedings of the Institution of Electrical Engineers,* 120(11), p.1321-1327.

CHAUHAN, B.C., DRANCE, S.M., DOUGLAS, G.R., JOHNSON, C.A. (1989). Visual field damage in normal-tension and high-tension glaucoma. *American Journal of Ophthalmology***,** 108(6), p.636-642

CHAUHAN, B.C., GARWAY-HEATH, D.F., GONI, F.J., ROSSETTI, L., BENGTSSON, B., VISWANATHAN, A.C., HEIJL, A. (2008). Practical recommendations for measuring rates of visual field change in glaucoma. *British Journal of Ophthalmology*, 92(4), p.569-573.

CHAUHAN, B.C., JOHNSON, C.A. (1999). Test-retest variability of frequency-doubling perimetry and conventional perimetry in glaucoma patients and normal subjects. *Investigative Ophthalmology and Visual Science,* 40, p.648-656

CHAUHAN, B.C., LEBLANC, R.P., SHAW, A.M., CHAN, A.B., MCCORMICK, T.A. (1997) Repeatable diffuse visual field loss in open-angle glaucoma. *Ophthalmology*, 104, p.532-538.

CHAUHAN, B.C., TOMPKINS, J.D., LEBLANC, R.P., McCORMICK, T.A. (1993). Characteristics of frequency-seeing curves in normal subjects, patients with suspected glaucoma and patients with glaucoma. *Investigative Ophthalmology and Visual Science*, 34, p.3534-3540.

CHEN, P.P. (2002). Correlation of visual field progression between eyes in patients with open-angle glaucoma. *Ophthalmology*, 109, p.2093-2099.

CHEN, P.P. and BUDENZ, D.L. (1998). The effects of cataract extraction on the visual field of eyes with chronic open-angle glaucoma. *American Journal of Ophthalmology*, 125(3), p.325-333.

CHUNG, H.J., CHOI, J.H., LEE, Y.C., KIM, S.Y. (2016). Effect of cataract opacity type and glaucoma severity on visual field index. *Optometry and Vision Science*, 93(6), p.575-578.

CHYLACK, L.T. Jr, WOLFE, J.K., SINGER, D.M., LESKE, M.C., BULLIMORE, M.A., BAILEY, I.L., FRIEND, J., MCCARTHY, D., WU, S.Y. (1993). The Lens Opacities Classification System III; the Longitudinal Study of Cataract Study Group. *Archives of Ophthalmology*, 111, p.831–836

COLLABORATIVE NORMAL-TENSION GLAUCOMA STUDY GROUP. (1998). Comparison of glaucomatous progression between untreated patients with normal-tension glaucoma and patients with therapeutically reduced intraocular pressures. *American Journal of Ophthalmology*, 126, p.487-497.

COLLABORATIVE NORMAL-TENSION GLAUCOMA STUDY GROUP. (1998a). The effectiveness of intraocular pressure reduction in the treatment of normal-tension glaucoma. *American Journal of Ophthalmology*, 126, p.498-505.

CONGDON, N., O'COLMAIN, B., KLAVER, C.C.,KLEIN, R., MUÑOZ, B., FRIEDMAN,D.S., KEMPEN, J., TAYLOR, H.R., MITCHELL, P.; Eye Diseases Prevalence Research Group (2004). Causes and prevalence of visual impairment among adults in the United States. *Archives of Ophthalmology*, 122, p.477-485.

CONWAY, M.L., HOSKING, S.L., ZHU, H., CUBBIDGE, R.P. (2014). Does the Swedish Interactive Threshold Algorithm (SITA) accurately map visual field loss attributed to vigabatrin? *BMC Ophthalmology*, 14, p.166.

CORALLO, G., ZINGIRIAN, M., GANDOLFO, E., CAPRIS, P., ROLANDO, M., FIORETTO, M. (1995). Updating the role of diffuse field loss in glaucoma diagnosis. IN MILLS, R.P., WALL, M. (eds): *Perimetry Update 1994/95*. Amsterdam/New York: Kugler Publications, p.283-287.

CUBBIDGE, R.P. (2005). *Visual Fields*. Edinburgh: Elsevier Butterworth Heinemann.

CUBBIDGE, R.P. (2012). Essentials of visual field assessment. Part 5: Analysis of visual fields data. *Optician*, 15th of June. Available from: **www.opticianonline.net** [Assessed 14January, 2016]

CURCIO, C.A. and ALLEN, K.A. (1990). Topography of ganglion cells in human retina. *Journal of Comparative Neurology*, 300, p.5-25.

DADA, T. and MANDAL, S. (2008). Short-wavelength automated perimetry. IN GARG, A., et al. (eds) *Techniques in glaucoma diagnosis and treatment.* New Delhi: Jaypee Brothers, p.87-91.

DE VOOGD, S., IKRAM, M.K., WOLFS, R.C.,(2005). Incidence of open-angle glaucoma in a general elderly population; the Rotterdam Study. *Ophthalmology*, 112, p.1487-1493.

DE WAARD, P.W., IJSPEERT, J.K., VAN DEN BERG, T.J., DE JONG, P.T. (1992). Intraocular light scattering in age-related cataracts. *Investigative Ophthalmology and Visual Science*, 33(3), p.618-625.

DE WIT, G.C., FRANSSEN, L., COPPENS, J.E., VAN DEN BERG, T.J. (2006). Simulating the straylight effects of cataracts. *Journal of Cataract and Refractive Surgery*, 32(2), p.294-300.

DELGADO, M.F., NGUYEN, N.T.A., COX, T.A., SINGH, K., LEE, D.A., DUEKER, D.K., FECHTNER, R.D., JUZYCH, M.S., LIN, S.C., NETLAND, P.A., PASTOR, S.A., SCHUMAN, J.S., SAMPLES, J.R. (2002). Automated Perimetry: a report by the American Academy of Ophthalmology. *Ophthalmology*, 109(12), p.2362-2374.

DEMIREL, S. (2015). Automated perimetry under the microscope: a re-examination of fundamental assumptions. *Investigative Ophthalmology & Visual Science*, 56, p.7224.

DEMIREL, S. and JOHNSON, C.A. (2001). Incidence and prevalence of short wavelength automated perimetry deficits in ocular hypertensive patients. *American Journal of Ophthalmology*, 131(6), p.709-715.

DENGLER-HARLES, M., WILD, J.M., COLE, M.D., O'NEILL, E.C., CREWS, S.J. (1990). The influence of forward light scatter on the visual field indices in glaucoma. *Graefe's Archive for Clinical and Experimental Ophthalmology*, 228(4), p.326-331.

DIAZ-ALEMAN, V.T., ANTON, A., DE LA ROSA, M.G., JOHNSON, Z.K., MCLEOD, S., AZUARA-BLANCO, A. (2009). Detection of visual-field deterioration by Glaucoma Progression Analysis and Threshold Noiseless Trend programs. *British Journal of Ophthalmology*, 93(3), p.322-328.

DIELEMANS I, VINGERLING, J.R., WOLFS, R.C.W., HOFMAN, A., GROBBEE, D.E., DE JONG, P.T.V.M. (1994) The prevalence of primary open-angle glaucoma in a population based study in the Netherlands. The Rotterdam Study. *Ophthalmology,* 101, p.1851-1855.

DONAHUE, S.P. and PORTER, A. (2001). SITA visual field testing in children. *Journal of American Association for Pediatric Ophthalmology and Strabismus*, 5 (2), p.114-117.

DONG, Z.M., WOLLSTEIN, G., SCHUMAN, J.S. (2016). Clinical utility of optical coherence tomography in glaucoma. *Investigative Ophthalmology and Visual Science*, 57, p.OCT556–OCT567

DRANCE, S., ANDERSON, D.R., SCHULZER, M. (2001). Risk factors for progression of visual field abnormalities in normal-tension glaucoma. *American Journal of Ophthalmology*,131, p.699-708

DRANCE, S.M. (1972). The glaucomatous visual field. *Investigative Ophthalmology and Visual Science*, 11(2), p.85-96.

DRANCE, S.M. (1991). Diffuse visual field loss in open-angle glaucoma. *Ophthalmology* 98(10), p.1533-1538.

DUGGAN, C., SOMMER, A., AUER, C., BURKHARD, K. (1985). Automated differential threshold perimetry for detecting glaucomatous visual field loss. *American Journal of Ophthalmology*, 100(3), p.420-423.

EDERER, F., GAASTERLAND, D.E., SULLIVAN, E.K., AGIS INVESTIGATORS (1994). The Advanced Glaucoma Intervention Study (AGIS) 1: Study design and methods and baseline characteristics of study patients. *Controlled Clinical Trials*, 15(4), p.299-325.

ELLIOT, D.B. (2007). *Clinical Procedures in Primary Eye Care. 3rd ed.* Edinburgh: Butterworth-Heinemann/Elsevier.

ELLIOTT DB, GILCHRIST J, WHITAKER D. (1989). Contrast sensitivity and glare sensitivity changes with three types of cataract morphology: are these techniques necessary in a clinical evaluation of cataract? *Ophthalmic and Physiological Optics,* 9, p.25-30.

ENOCH, J. (1978). Quantitative layer-by-layer perimetry. *Investigative Ophthalmology and Visual Science*, 17, p.208-257.

FABRIKANTOV, O.L., SHUTOVA, S.V., SUKHORUKOVA, A.V. (2015). Comparative characteristics of the standard automated perimetry and contour perimetry methods in diagnosis the initial stage of glaucoma. *The Fyodorov Journal of Ophthalmic Surgery*, 4, p.24-29.

FANIHAGH, F., KREMMER, S., ANASTASSIOU, G., SCHALLENBERG, M. (2015). Optical coherence tomography, scanning laser polarimetry and confocal scanning laser ophthalmoscopy in retinal nerve fiber layer measurements of glaucoma patients. *The Open Ophthalmology Journal*, 31(9), p.41-48.

FANKAUSER, F. (1982). Developmental milestones of automated perimetry. IN: HENDKIND, P. (ed.) *ACTA: XXIV International Congress of Ophthalmology*. Philadelphia, PA: JB Lippincott; p.147-150.

FANKHAUSER, F. and HAEBERLIN, H. (1980). Dynamic range and stray light. An estimate of the falsifying effects of stray light in perimetry. *Documenta Ophthalmologica*, 50, p.143-167.

FANKHAUSER, F., KOCH, P., ROULIER, A. (1972). On automation of perimetry. *Albrecht von Graefe's Archive for Clinical and Experimental Ophthalmology*, 184, p.126-150.

FANKHAUSER, F., SPAHR, J., BEBIE, H. (1977). Three years of experience with the 'Octopus' automatic perimeter. *Documenta Ophthalmologica Proc Series*, 14, p.7-5.

FAUL, F., ERDFELDER, E., LANG, A. G., & BUCHNER, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39(2), p.175-191.

FECHTNER, R..D. and WEINREB, R.N. (1994). Mechanisms of optic nerve damage in primary open angle glaucoma. *Survey of Ophthalmology*, 39, p.23-42.

FINGERET, M. (2016). Do we need perimetry? *Glaucoma Today*, May/June, p.32-34.

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

FLAMMER, J., DRANCE, S.M. and SCHULTZER, M. (1984a). Covariates of long-term fluctuation of differential threshold. *Archives of Ophthalmology*, 102(6), p.880-882.

FLAMMER, J., DRANCE, S.M., AUGUSTINY, L., FUNKHOUSER, A. (1985). Quantification of glaucomatous visual field defects with automated perimetry. *Investigative Ophthalmology and Visual Science*, 26, p.176-181.

FLAMMER, J., DRANCE, S.M., ZULAUF, M. (1984). Differential light threshold. Short- and long-term fluctuation in patients with glaucoma, normal controls, and patients with suspected glaucoma. *Archives of Ophthalmology*, 102(5), p.704-706.

FLANAGAN, J.G. (2009). Visual field examination. IN: ROSENFIELD, M., LOGAN, N., EDWARDS, K.H. (eds). *Optometry: Science, Techniques and Clinical Management*. Edinburgh: Elsevier Health Sciences, p.317-334.

FLANAGAN, J.G., MOSS, I.D., WILD, J.M., HUDSON, C., PROKOPICH, L., WHITAKER, D., O'NEILL, E.C. (1993a). Evaluation of FASTPAC: a new strategy for threshold estimation with the Humphrey Field Analyser. *Graefe's Archive for Clinical and Experimental Ophthalmology*, 231(8), p.465-469.

FLANAGAN, J.G., WILD, J., TROPE, G. (1993). Evaluation of FASTPAC, a new strategy for thresholds estimation with Humphrey field analyser, in a glaucomatous population. *Ophthalmology*, 100(6), p.949-954.

FLANAGAN, J.G., WILD, J., TROPE, G. (1993b). The visual field indices in primary open-angle glaucoma. *Investigative Ophthalmology and Visual Science*, 34, p.2266-2274.

FLAXMAN, S.R., BOURNE, R.R.A., RESNIKOFF, S., ACKLAND, P., BRAITHWAITE, T., CICINELLI, M.V., DAS, A., JONAS, J.B., KEEFFE, J., KEMPEN, J.H., LEASHER, J., LIMBURG, H., NAIDOO, K., PESUDOVS, K., SILVESTER, A., STEVENS, G.A., TAHHAN, N., WONG, T.Y., TAYLOR, H.R., on behalf of the Vision Loss Expert Group of the Global Burden of Disease Study. (2017). Global causes of blindness and distance vision impairment 1990–2020: a systematic review and meta-analysis. *Lancet Glob Health,* 5, p.e1221-34

FOGAGNOLO, P., ROSSETTI, L., RANNO, S., FERRERAS, A., ORZALESI, N. (2008). Short-wavelength automated perimetry and frequency-doubling technology perimetry in glaucoma. *Progress in Brain Research*, 173, p.101-124.

FORSMAN E, KIVELA T, VESTI E. (2007). Lifetime visual disability in open-angle glaucoma and ocular hypertension. *Journal of Glaucoma*, 16, p.313-319.

FREDETTE, M.J., GIGUÈRE, A., ANDERSON, D.R**.**, BUDENZ, D.L., MCSOLEY, J. (2015). Comparison of Matrix with Humphrey Field Analyzer II with SITA. *Optometry and Vision Science*, 92(5), p.527-536.

FREEMAN, E.E., MUNOZ, B., RUBIN, G., WEST, S.K. (2007). Visual field loss increases the risk of falls in older adults: The Salisbury eye evaluation. *Investigative Ophthalmology and Visual Science*, 48, p.4445-4450.

FRIEDMAN, D.S., WOLFS, R.C., O'COLMAIN, B.J., KLEIN, B.E., TAYLOR, H.R., WEST, S., LESKE, M.C., MITCHELL, P., CONGDON, N., KEMPEN, J.; Eye Diseases Prevalence Research Group. (2004). Prevalence of open-angle glaucoma among adults in the United States. *Archives of Ophthalmology*, 122(4), p.532-538.

FUJIMOTO, N, MINOWA, K., MIYAUCHI, O, HANAWA, T., ADACHI-USAMI, E., (2002). Learning effect for frequency doubling perimetry in patients with glaucoma. *American Journal of Ophthalmology*, 133(2), p.269-270.

FUNKHOUSER, A. and FANKHAUSER, F. (1991). The effects of weighting the “mean defect” visual field index according to threshold variability in thecentral and midperipheral visual field. *Graefe’s Archive for Clinical and Experimental Ophthalmology,* 229, p.228-231.

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

GARDINER, S.K., DEMIREL, S. and JOHNSON, C.A. (2008). Is there evidence for continued learning over multiple years in clinical perimetry? *Optometry and Vision Science*, 85(11), p.1043-1048.
GARDINER, S.K., DEMIREL, S., GOREN, D., MANSBERGER, S.L., SWANSON, W.H. (2015). The effect of stimulus size on the reliable stimulus range of perimetry. *Translational Vision Science and Technology*, 4(2), p.10.

GARDINER, S.K., JOHNSON, C.A., SPRY, P.G. (2006). Normal age-related sensitivity loss for a variety of visual functions throughout the visual field. *Optometry and Vision Science*, 83, p.438-443.

GARDINER, S.K., SWANSON, W.H., GOREN, D., MANSBERGER, S.L., DEMIREL, S. (2014). Assessment of the reliability of standard automated perimetry in regions of glaucomatous damage. *Ophthalmology*, 121, p.1359-1369.

GARDINER, S.K., JOHNSON, C.A., DEMIREL, S. (2012). The effect of test variability on the structure-function relationship in early glaucoma. *Graefe's Archive for Clinical and Experimental Ophthalmology*, 250(12), p.1851-1861.

GARG, A. (2008). Visual fields. IN GARG, A., et al. (eds) *Techniques in glaucoma diagnosis and treatment.* New Delhi: Jaypee Brothers, p.78-86

GARWAY-HEATH, D.F., POINOOSAWMY, D., FITZKE, F.W., HITCHINGS, R.A. (2000). Mapping the visual field to the optic disc in normal tension glaucoma eye. *Ophthalmology*, 107, p.1809-1815.

GAZZARD, FOSTER, P.J., DEVEREUX, J.G., OEN, F., CHEW, P., KHAW, P.T., SEAH, S. (2003). Intraocular pressure and visual field loss in primary angle closure and primary open angle glaucomas. *British Journal of Ophthalmology*, 87, p.720-725.

GAZZARD, G., FOSTER, P.J., VISWANATHAN, A.C., DEVEREUX, J.G., OEN, F.T.S., CHEW, P.T.K., KHAW, P.T., SEAH, S.K.L. (2002). The severity and spatial distribution of visual field defects in primary glaucoma: a comparison of primary open-angle glaucoma and primary angle-closure glaucoma. *Archives of Ophthalmology*, 120, p.1636-1643.

GHAZALI, N., ASLAM, T., HENSON, D.B. (2015). New superior-inferior visual field asymmetry indices for detecting POAG and their agreement with glaucoma hemifield test. *Eye*, 29, p.1375-1382.

GILLESPIE, B.W., MUSCH, D.C., GUIRE, K.E., MILLS, R.P., LICHTER, P.R., JANZ, N.K., WREN, P.A. (2003). The collaborative initial glaucoma treatment study: baseline visual field and test-retest variability. *Investigative Ophthalmology and Visual Science,* 44(6), p.2613-2620.

GILLIES, W.E. and BROOKS, A.M. (1998). Effect of lens opacity on the glaucomatous field of vision. *Australian New Zealand Journal of Ophthalmology*, 26 (Suppl), p.19-21

GIRKIN, C.A., EMDADI, A., SAMPLE, P.A., BLUMENTHAL, E.Z., LEE, A.C., ZANGWILL, L.M., WEINREB, R.N. (2000). Short-wavelength automated perimetry and standard perimetry in the detection of progressive optic disc cupping. *Archives of Ophthalmology*, 118(9), p.1131-1236.

GLASS, E., SCHAUMBERGER, M., LACHENMAYR, B.J. (1995). Simulations for Fastpac and the standard 4-2 dB Full-Threshold strategy of Humphrey Field Analyser. *Investigative Ophthalmology and Visual Science*, 36(9), p.1847–1854.

GLEN, F.C., BAKER, H., CRABB, D.P. (2014). A qualitative investigation into patients' views on visual field testing for glaucoma monitoring. *BMJ Open*.4:e003996.

GLIKLICH, R.E., STEINMANN, W.C., SPAETH, G.L. (1989). Visual field change in low-tension glaucoma over a five-year follow-up. *Ophthalmology*, 96, p.316-320.

GLOOR, B., SCHMIED, U., FAESSLER, A. (1980). Changes of glaucomatous field defects. Degree of accuracy of measurements with the automatic perimeter Octopus. *International Ophthalmology*, 3, p.5-10.

GLOOR, B.P. (2009). Franz Fankhauser: the father of automated perimeter. *Survey of Ophthalmology*, 54, p.417-425.

GLOOR, B.P., SCHMIED, U., FÄSSLER, A. (1981) Changes of glaucomatous field defects analysis of OCTOPUS fields with programme Delta. IN: GREVE, E.L., VERRIEST, G. (eds) Fourth International Visual Field Symposium Bristol, April 13–16, 1980. Documenta Ophthalmologica Proceedings Series, vol 26. Springer, Dordrecht

GOLDBAUM, M.H., SAMPLE, P.A., CHAN, K., WILLIAM, J., LEE, T.W., BLUMENTHAL, E., GIRKIN, C.A., ZANGWILL, L.M., BOWD, C., SEJNOWSKI, T., WEINREB, R.N. (2002). Comparing machine learning classifiers for diagnosing glaucoma from standard automated perimetry. *Investigative Ophthalmology and Visual Science*, 43, p.162-169.

GOLDBAUM, M.H., SAMPLE, P.A., ZHANG, Z., CHAN, K., HAO, J., LEE, T., BODEN, C., BOWD, C., BOURNE, R., ZANGWILL, L., SEJNOWSKI, T., SPINAK, D., WEINREB, R.N. (2005). Using unsupervised learning with independent component analysis to identify patterns of glaucomatous visual field defects. *Investigative Ophthalmology and Visual Science,* 46, p. 3676-3683.

GONZALEZ DE LA ROSA M, GONZALEZ-HERNANDEZ M, SANCHEZ-MENDEZ M, MEDINA-MESA E, RODRIGUEZ DE LA VEGA R.(2010). Detection of morphological and functional progression in initial glaucoma. *British Journal of Ophthalmology*, 94(4), p.414-418.

GONZALEZ DE LA ROSA, M. and GONZALEZ-HERNANDEZ, M. (2011). Monitoring visual field progression. *British Journal of Ophthalmology*, 95 (2), p.157-158.

GONZALEZ DE LA ROSA, M. and GONZALEZ-HERNANDEZ, M. (2013). A strategy for averaged estimates of visual field threshold: Spark. *Journal of Glaucoma*, 22, p.284-289.

GONZALEZ DE LA ROSA, M. and PAREJA, A. (1997). Influence of the “fatigue effect” on the mean deviation measurement in perimetry. *European Journal of Ophthalmology*, 7, p.29-34.

GONZALEZ DE LA ROSA, M., ARMAS-DOMINGUEZ, K., DIAZ-ALEMAN, T., GONZALEZ-HERNANDEZ, M., JEREZ-FIDALGO, M. (2010a). Specificity of the program Threshold Noiseless Trend (TNT) for perimetric progression analysis. *Current Eye Research*, 35(4), p.302-307.

GONZALEZ DE LA ROSA, M., GONZALEZ HERNANDEZ, M., ABRALDES, M., AZUARA-BLANCO, A. (2002a). Quantification of inter-point topographic correlations of threshold values in glaucomatous visual fields. *Journal of Glaucoma*, 11, p.30-34.

GONZALEZ DE LA ROSA, M., GONZALEZ HERNANDEZ, M., AGUILAR ESTEVEZ, J., ABREU REYES, A., PAREJA RÍOS, A. (2002). Classificación topográfica del campo visual glaucomatoso*. Archivos de la Sociedad Española de Oftalmología*, 77, p.87-94.

GONZALEZ DE LA ROSA, M., GONZALEZ-HERNANDEZ, M., SANCHEZ-GARCIA, M., RODRIGUEZ DE LA VEGA, R., DIAZ-ALEMAN, T., RIOS, A.P. (2013). Oculus-Spark perimetry compared with 3 procedures of glaucoma morphologic analysis (GDx, HRT, and OCT). *European Journal of Ophthalmology*, 23(3), p.316–323.

GONZALEZ DE LA ROSA, M., MARTINEZ, A., SANCHEZ, M., MESA, C., CORDOVES, L., LOSADA, M.J. (1997). Accuracy of the Tendency-Oriented Perimetry (TOP) in the Octopus 1-2-3 Perimeter. IN: WALL, M. and WILD, J.M. (eds.). *Perimetry Update 1996/97.*Amsterdam/ New York: Kugler Publication., p.119-223.

GONZALEZ DE LA ROSA, M., GONZALEZ-HERNANDEZ, M., ALAYON, S. (2015). Glaucoma morphologic damage estimated from functional tests. *European Journal of Ophthalmology*, 25(6), p.496-502

GONZALEZ DE LA ROSA, M., GONZALEZ-HERNANDEZ, M., SIGUT, J., ALAYON, S., RADCLIFFE, N., MENDEZ-HERNANDEZ, C., GARCÍA-FEIJOO, J., FUERTES-LAZARO, I., PEREZ-OLIVAN, S.,FERRERAS, A. (2013a). Measuring hemoglobin levels in the optic nerve head: comparisons with other structural and functional parameters of glaucoma. *Investigative Ophthalmology and Visual Science*, 54(1), p.482-489.

GONZALEZ DE LA ROZA, M. (2014). Oculus SPARK: a new standard in visual field examination. Available from: https**://www.genop.co.za/media/pdf/Oculus/Oculus%20SPARK%20Brochure.pdf.** [Assessed 20 January 2017]

GORDON, M.O. and KASS, M.A. (1999). The Ocular Hypertension Treatment Study: design and baseline description of the participant. *Archives of Ophthalmology*, 117(5), p.573-583.

GORDON, M.O., BEISER, J.A., BRANDT, J.D., (2002). The Ocular Hypertension Treatment Study: baseline factors that predict the onset of primary open-angle glaucoma. *Archives of Ophthalmology*, 120, p.714-720; discussion, p.829-830.

GRAMER, E., GERLACH, R., KRIEGLSTEIN, G.K., LEYDHECKER, W. (1982). Topography of early glaucomatous visual field defects in computerized perimetry. *Klinische Monatsblätter für Augenheilkunde,* 180, p.515-523.

GREVE, E.L. (1973). *Single and multiple stimulus static perimetry: The two phases of perimetry*. The Hague: Dr. W. Junk Publishers.

GREWE, R. (1986). The history of glaucoma [Article in German].*Klinische Monatsblätter für Augenheilkunde*, 188(2), p.167-169.

GUTHAUSER, U. and FLAMMER, J. (1988). Quantifying visual field damage caused by cataract. *American Journal of Ophthalmology*, 106(4), p.480-484.

GUTHAUSER, U., FLAMMER, J., NIESEL, P. (1987). Relationship between cataract density and visual field damage. In: GREVE, E.L. and HEIJL, A. (eds) *Seventh International Visual Field Symposium.* Doc Ophthalmol Proc Ser 49. Dordrecht:Martinus Nijhoff/Dr W Junk, p.39-41.

HAAS, A., FLAMMER, J., SCHNEIDER, U. (1986). Influence of age on the visual fields of normal subjects. *American Journal of Ophthalmology*, 101, p.199-203.

HALLET, P.E. (1969). The variations in visual threshold measurement. *Journal of Physiology*, 202, p.403-419.

HART, W.M.JR. and BECKER B. (1982). The onset and evolution of glaucomatous visual field defects. *Ophthalmology,* 89, p.268-279.

HARVEY, B. (2011). Crossing the threshold. *Optician*. Feb, 11, 2011. Available from: **https://s3-eu-west-1.amazonaws.com/rbi-communities/wp-content/uploads/importedimages/henson.pdf.** [Assesed 28June, 2015]

HARWERTH, R.S., VILUPURU, A.S., RANGASWAMY, N.V., SMITH, E.L. 3rd. (2007). The relationship between nerve fiber layer and perimetry measurements. *Investigative Ophthalmology and Visual Science*, 48(2), p.763-773.

HATTENHAUER, M.G., JOHNSON, D.H., ING, H.H., et al. (1998). The probability of blindness from open angle glaucoma. *Ophthalmology,* 105, p.2099-2104.

HAYASHI, K., HAYASHI, H., NAKAO, F., HAYASHI, F. (2001). Influence of cataract surgery on automated perimetry in patients with glaucoma. *American Journal of Ophthalmology*, 132(1), p.41-46.

HEEG, G.P., PONSIOEN, T.L., JANSONIUS, N.M. (2003). Learning effect, normal range, and test–retest variability of Frequency Doubling Perimetry as a function of age, perimetric experience, and the presence or absence of glaucoma. *Ophthalmic and Physiological Optics*, 23(6), p.535-540

HEIDER, H.W., SEEZ, K.J., SCHNAUDIGEL, O.E. (1991). Changes in the visual field caused by lens opacities.*Klinische Monatsblätter für Augenheilkunde*, 198, p.15-19.

HEIJL A. (1977a). Time changes of contrast thresholds during automatic perimetry. *Acta Ophthalmologica,* 55(4), p.696-708.

HEIJL, A. (1977). Computer test logics for automated perimetry. *Acta Ophthalmologica (Copenhagen)*, 55, p.15-31.

HEIJL, A. (1989). Lack of diffuse loss of differential light sensitivity in early glaucoma. *Acta Ophthalmologica* (Copenh), 67, p.353-360.

HEIJL, A. (1989a). Visual field changes in early glaucoma and how to recognize them. *Survey of Ophthalmology*, 33(suppl), p.403-404.

HEIJL, A. and DRANCE, S.M. (1983). Changes in differential threshold in patients with glaucoma during prolonged perimetry. *British Journal of Ophthalmology*, 67(8), p.512-516.

HEIJL, A. and KRAKAU, C.E.T. (1975). An automatic static perimeter, design and pilot study. *Acta Ophthalmologica*, 53(3), p.293-310.

HEIJL, A. and LUNDQVIST, L. (1984). The frequency distribution of earliest glaucomatous visual field defects documented by automatic perimetry. *Acta Ophthalmologica (Copenhagen),* 62, p.658-664.

HEIJL, A. and PATELLA, V.M. (2002). *Essential Perimetry. The Field Analyzer Primer*(3rd ed), Dublin California: Carl Zeiss Meditec,

HEIJL, A., BENGTSSON, B. (1996). The effect of perimetric experience in patients with glaucoma. *Archives of Ophthalmology*, 14, p.19-22.

HEIJL, A., BENGTSSON, B., CHAUHAN, B.C., LIEBERMAN, M.F., CUNLIFFE, I., HYMAN, L., LESKE, M.C. (2008). A comparison of visual field progression criteria of 3 major glaucoma trials in early manifest glaucoma trial patients. *Ophthalmology* 115:1557-1565

HEIJL, A., BENGTSSON, B., PATELLA, V.M. (2000). Glaucoma follow-up when converting from long to short perimetric threshold tests. *Archives of Ophthalmology*, 118, p.489-493.

HEIJL, A., LESKE, M.C., BENGTSSON, B., BENGTSSON, B., HUSSEIN, M. (2003). Early Manifest Glaucoma Trial Group. Measuring visual field progression in the Early Manifest Glaucoma Trial. *Acta Ophthalmologica Scandinavica*, 81, p.286-293.

HEIJL, A., LINDGREN, A., LINDGREN, G. (1989a). Test-retest variability in glaucomatous visual fields. *American Journal of Ophthalmology*, 108, p.130-135.

HEIJL, A., LINDGREN, G., OLSSON, J. (1987). Normal variability of static perimetric threshold values across the central visual field. *Archives of Ophthalmology*, 105, p.1544-1549.

HEIJL, A., LINDGREN, G., OLSSON, J. (1987a). A package for the statistical analysis of visual field. *Documenta Ophthalmologica Proceedings Series*, 49, p.153-168.

HEIJL, A., LINDGREN, G., OLSSON, J. (1989). The effect of perimetric experience in normal subjects. *Archives of Ophthalmology,* 107, p.81-86.

HEIJL, A., LINDGREN, G., OLSSON, J., ASMAN, P. (1992). On weighted visual field indices. *Graefe's Archive for Clinical and Experimental Ophthalmology,* 230, p.397-400.

HEIJL, A., PATELLA, V.M., BENGTSSON, B. (2012). *The Field Analyzer Primer: Effective Perimetry*. 4th ed. Dublin, CA: Carl Zeiss Meditec, Inc.

HEIJL, A, PATELLA, V.M., CHONG, L.X., IWASE, A., LEUNG, C.K., TUULONEN, A., LEE, G.C., CALLAN, T., BENGTSSON, B. (2019). A new SITA perimetric threshold testing algorithm; construction and a multi-center clinical study. *American Journal of Ophthalmology*, 198, p.154-165.

HEIJL, A., LINDGREN, G., OLSSON, J. (1988). Perimetric threshold variability and age. *Archives of Ophthalmology*, 106(4), p.450-452.

HENSON, D.B. and CHAUHAN, B.C. (1985). Informational content of visual field location in glaucoma. *Documenta Ophthalmologica,* 59, p.341-352.

HENSON, D.B. and EMUH, T. (2010). Monitoring vigilance during perimetry with pupillography. *Investigative Ophthalmology and Visual Science*, 51, p.3540-3543.

HENSON, D.B., ARTES, P.H., CHAUHAN, B.C. (1999). Diffuse loss of sensitivity in early glaucoma. *Investigative Ophthalmology and Visual Science*, 40, p.3147-3151.

HENSON, D.B., BRYSON, H. (1991). Is the variability in glaucomatous field loss due to poor fixation control? In: MILLS RP, HEIJL A, eds. Perimetry Update 1990/1991. Amsterdam: Kugler & Ghendini; p.217-220

HENSON, D.B., CHAUDRY, S., ARTES, P.H., FARAGHER, E.B., ANSONS, A. (2000). Response variability in the visual field: comparison of optic neuritis, glaucoma, ocular hypertension and normal eyes. *Investigative Ophthalmology and Visual Science*, 41(2), p.417-421.

HENSON, D.B., CHAUHAN, B.C., HOBLEY, A. (1988). Screening for glaucomatous visual field defects: the relationship between sensitivity, specificity and the number of test locations. *Ophthalmic and Physiological Optics*, 8, p.123-127.

HENSON, D.B., DIX, S.M., OBORNE, A.C. (1984). Evaluation of the Friedmann Visual Field Analyser Mark II. Part 1. Results from a normal population. *British Journal of Ophthalmology*, 68, p.458-462.

HENSON, D.B., EVANS, J., CHAUHAN, B.C. (1996). Influence of fixation accuracy on threshold variability in patients with open angle glaucoma. *Investigative Ophthalmology and Visual Science*, 37, p.444-450.

HENSON, D.B., EVANS, J., LANE, C.M. (1995). Fixation accuracy of patients with glaucoma during full threshold perimetry. In: MILLS, R.P., WALL, M., eds. *Perimetric Update 1994/95.* Amsterdam: Kugler; p.241-248

HERMANN, A., PAETZOLD, J., VONTHEIN, R., KRAPP, E., RAUSCHER, S., SCHIEFER, U. (2008). Age-dependent normative values for differential luminance sensitivity in automated static perimetry using the Octopus 101. *Acta Ophthalmologica Scandinavica*, 86(4), p.446-455.

HEUER, D.K., ANDERSON, D.R., KNIGHTON, R.W., FEUER, W.J., GRESSEL, M.G. (1988). The influence of simulated light scattering on automated perimetric threshold measurements. *Archives of Ophthalmology*, 106(9), p.1247-1251.

HIRASAWA, K. and SHOJI, N. (2016). Swedish Interactive Threshold Algorithm for central visual field defects unrelated to nerve fiber layer. *Graefe's Archive for Clinical and Experimental Ophthalmology*, 254(5), p.845-854.

HIRASAWA, K., SHOJI, N., KASAHARA, M., MATSUMURA, K., SHIMIZU, K. (2016). Comparison of size modulation and conventional standard automated perimetry with the 24-2 test protocol in glaucoma patients. *Scientific Reports*, May 5; 6: 25563.

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

HOFFMANN, E.M., BODEN, C., ZANGWILL, L.M., BOURNE, R.R., WEINREB, R.N., SAMPLE, P.A. (2006). Inter-eye comparison of patterns of visual field loss in patients with glaucomatous optic neuropathy. *American Journal of Ophthalmology,* 141(4), p.703-708.

HOLMIN, C. and KRAKAU, C.E.T. (1979). Variability of glaucomatous visual field defects in computerized perimetry. *Graefe's Achivefor Clinical and Experimental Ophthalmology*, 210, p.235-250.

HONG, S., AHN, H., HA, S.J., YEOM, H.Y., SEONG, G.J., HONG, Y.J. (2007). Early glaucoma detection using Humphrey Matrix Perimeter, GDx VCC, Stratus OCT, and retinal nerve fiber layer photography. *Ophthalmology*, 114(2), p.210-215.

HOOD, D.C., RAZA, A.S., DE MORAES, C.G.V., ODEL, J.G., GREENSTEIN, V.C., LIEBMANN, J.M., RITCH, R. (2011). Initial arcuate defects within the central 10 degrees in glaucoma. *Investigative Ophthalmology and Visual Science*, 52(2), p.940-946.

HORANI, A., FRENKEL, S., YAHALOM, C., FARBER, M.D., TICHO, U., BLUMENTHAL, E.Z. (2002). The learning effect in visual field testing of healthy subjects using frequency doubling technology *Journal of Glaucoma*, 11(6), p.511-516

HORN, F.K., MARDIN, C.Y., VIESTENZ, A., JÜNEMANN, A.G. (2005). Association between localized visual field losses and thickness deviation of the nerve fiber layer in glaucoma. *Journal of Glaucoma*, 14, p.419-425.

HUDSON, C., WILD, J.M., and O'NEILL, E.C. (1994). Fatigue effect during a single session of automated static threshold perimetry*. Investigative Ophthalmology & Visual Science*, 35, p.268–280.

IMAGING AND PERIMETRY SOCIETY (2012). Imaging and Perimetry Society (IPS) : What is Perimetry?. Available from: **http://www.perimetry.org/Perimetr.htm** [Assessed 21Feb, 2015]

IMAGING AND PERIMETRY SOCIETY (2012a). Imaging and Perimetry Society (IPS): History of the visual field before perimetry. Available from: **http://www.perimetry.org/PerimetryHistory/1-pre-perimetry.htm** [Assessed 2April, 2015]

INAZUMI, K., TSUJI, A., YAMAMOTO, T. and KITAZAWA, Y. (1998). Evaluation of the Swedish Interactive Thresholding Algorithm, a new thresholding algorithm, of the Humphrey Field Analyzer in glaucoma patients. *Nippon GankaGakkaiZasshi*, 102 (10), p.667-672. [Japanese]

ISHIYAMA, Y., MURATA, H., ASAOKA, R. (2015). The usefulness of gaze tracking as an index of visual field reliability in glaucoma patients. *Investigative Ophthalmology and Visual Science*, 56, p. 6233-6236.

ISHIYAMA, Y., MURATA, H., MAYAMA, C., ASAOKA, R. (2014). An objective evaluation of gaze tracking in Humphrey perimetry and the relation with the reproducibility of visual fields: a pilot study in glaucoma. *Investigative Ophthalmology and Visual Science*, 55(12), p.8149–8152

IWASE, A., KITAZAWA, Y., OHNO, Y. (1988). On age-related norms of the visual field. *Japanese Journal of Ophthalmology*, 32(4), p.429-437.

JAFFE, G.J., ALVARADO, J.A., JUSTER, R.P. (1986). Age-related changes of the normal visual field. *Archives of Ophthalmology*, 104(7), p.1021-1025.

JAMPEL, H., VITALE, S., DING, Y., QUIGLEY, H., FRIEDMAN, D., CONGDON, N., ZEIMER, R.. (2006). Test-retest variability in structural and functional parameters of glaucoma damage in the glaucoma imaging longitudinal study. *Journal of Glaucoma*, 15(2), p.152-157.

JAMPEL, H.D., SINGH, K., LIN, S.C., CHEN, T.C., FRANCIS, B.A., HODAPP, E., SAMPLES, J.R., SMITH, S.D. (2011). Assessment of visual function in glaucoma: A report by the American Academy of Ophthalmology. *Ophthalmology,* 118(5), p.986-1001.

JANOSIK, E. and MARCZAK, W. (2016). The effect of warm and cool lighting on visual performance of elderly workers. *Zeszyty Naukowe Politechniki Poznańskiej. Organizacja i Zarządzanie*, Nr 70, p.51-67.

JOHNSON, C.A. (2013). Psychophysical factors that have been applied to clinical perimetry. *Vision Research*, 90, p.25-31.

JOHNSON, C.A. and SAMUELS, S.J. (1997). Screening for glaucomatous visual field loss with frequency-doubling perimetry. *Investigative Ophthalmology and Visual Science*, 38, p.413-425.

JOHNSON, C.A. and NELSON-QUIGG, M. (1993). A prospective three-year study of response properties of normal subjects and patients during automated perimetry. *Ophthalmology*, 100, p.269–274.

JOHNSON, C.A. and SAMUELS, S.J. (1997). Screening for glaucomatous visual field loss using the frequency-doubling contrast test. *Investigative Ophthalmology and Visual Science*, 38, p.413-425.

JOHNSON, C.A., ADAMS, A.J., CASSON, E.J., BRANDT, J.D. (1993b). Progression of early glaucomatous visual field loss as detected by blue-on-yellow and standard white-on-white automated perimetry. *Archives of Ophthalmology*, 111, p.651-656.

JOHNSON, C.A., ADAMS, A.J., CASSON, E.J., BRANDT,J.D. (1993a). Blue-on-yellow perimetry can predict the development of glaucomatous visual field loss. *Archives of Ophthalmology*, 111(5), p.645-650.

JOHNSON, C.A., ADAMS, C.W., LEWIS, R.A. (1988). Fatigue effects in automated perimetry. *Applied Optics*, 27, p.1030-1037.

JOHNSON, C.A., CHAUHAN, B.C., SHAPIRO, L.R. (1993). Properties of staircase procedures for estimating thresholds in automated perimetry. *Investigative Ophthalmology and Visual Science,* 33, p.2966-2974.

JOHNSON, C.A., KELTNER, J.L. and BALESTRERY, F.G. (1981). Static and acuity profile perimetry at various adaptation levels. *Documenta Ophthalmologica*, 50, p.371-388.

JOHNSON, C.A., KELTNER, J.L., CELLO, K.E., EDWARDS, M., KASS, M.A., GORDON, M.O., BUDENZ, D.L., GAASTERLAND, D.E., WERNER, E. (2002). Ocular Hypertension Study Group. Baseline visual field characteristics in the Ocular Hypertension Treatment Study. *Ophthalmology*, 109 (3), p.432-437.

JOHNSON, C.A., SAMPLE P.A., CIOFFI, G.A., LIEBMANN, J.R., WEINREB, R.N. (2002a). Structure and function evaluation (SAFE): I. criteria for glaucomatous visual field loss using standard automated perimetry (SAP) and short wavelength automated perimetry (SWAP). *American Journal of Ophthalmology*, 134, p.177-185.

JOHNSON, C.A., WALL, M., THOMPSON, H.S. (2011). A history of perimetry and visual field testing. *Optometry and Vision Science*, 88, p.E8-E15.

JOSON, P.J., KAMANTIGUE, M.E.G., CHEN, P.P. (2002). Learning effects among perimetric novices in frequency doubling technology perimetry. *Ophthalmology*, 109(4), p.757-760.

JUNOY MONTOLIO, F.G., WESSELINK, C., GORDIJN, M., 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 and Visual Science*, 53, p.7010-7017

KANAMORI, A., NAGAI-KUSUHARA, A., ESCAÑO, M.F., MAEDA, H., NAKAMURA, M., NEGI, A.(2006). Comparison of confocal scanning laser ophthalmoscopy, scanning laser polarimetry and optical coherence tomography to discriminate ocular hypertension and glaucoma at an early stage. *Graefe's Archive for Clinical and Experimental Ophthalmology*,244(1), p.58-68..

KANG, E.M., HONG, S., KIM, C.Y., SEONG, G.J. (2015). Relationship between peripapillary retinal nerve fiber layer thickness measured by optical coherence tomography and visual field severity indices. *Korean Journal of Ophthalmology*, 29(4), p.263-269.

KAPETANAKIS, V.V**.**, CHAN, M.P**.**, FOSTER, P.J**.**, COOK, D.G**.**, OWEN, C.G**.**, RUDNICKA, A.R**.** (2016). Global variations and time trends in the prevalence of primary open angle glaucoma (POAG): a systematic review and meta-analysis. *British Journal of Ophthalmology,* 100(1), p.86-93

KARBASSI, M., KHU, P.M., SINGER, D.M., CHYLACK, L.T. Jr. (1993). Evaluation of Lens Opacities Classification System III applied at the slitlamp. *Optometry and Vision Science*, 70, p.923–928

KATZ J, GILBERT D, QUIGLEY HA, SOMMER A. (1997). Estimating progression of visual field loss in glaucoma. *Ophthalmology*, 104(6), p.1017-1025.

KATZ J, QUIGLEY HA, SOMMER A. (1996). Detection of incident field loss using the glaucoma hemifield test. *Ophthalmology*, 103, p.657-663.

KATZ J. (1999). Scoring systems for measuring progression of visual field loss in clinicaltrials of glaucoma treatment. *Ophthalmology*, 106(2), p. 391 – 395.

KATZ, J. (2000). A comparison of the pattern- and total deviation-based glaucoma change probability programs. *Investigative Ophthalmology and Visual Science*, 41, p.1012-1016.

KATZ, J. and SOMMER, A. (1987). A longitudinal study of the age-adjusted variability of automated visual fields. *Archives of Ophthalmology,* 105, p.1083-1086.

KATZ, J. and SOMMER, A. (1988). Reliability indexes of automated perimetric tests. *Archives of Ophthalmology*, 106, p.1252-1254.

KATZ, J. and SOMMER, A. (1990). Screening for glaucomatous visual field loss. The effect of patient reliability. *Ophthalmology,* 97, p.1032-1037.

KATZ, J., CONGDON, N., FRIEDMAN, D.S. (1999). Methodological variations in estimating apparent progressive visual field loss in clinical trials of glaucoma treatment. *Archives of Ophthalmology*, 117(9), p.1137-1142.

KATZ, J., QUIGLEY, H.A., SOMMER, A. (1995a). Repeatability of the glaucoma hemifield test in automated perimetry. *Investigative Ophthalmology and Visual Science*, 36(8), p.1658-1664.

KATZ, J., QUIGLEY, H.A., SOMMER, A. (1996). Detection of incident field loss using the glaucoma hemifield test. *Ophthalmology*, 103, p.657-63

KATZ, J., SOMMER, A., GAASTERLAND, D.E., ANDERSON, D.R. (1991). Comparison of analytic algorithms for detecting glaucomatous visual field loss. *Archives of Ophthalmology*, 109(12), p.1684-1689.

KATZ, J., SOMMER, A., WITT, K. (1991a). Reliability of visual field results over repeated testing. *Ophthalmology*, 98, p.70-75.

KATZ, J., TIELSCH, J.M., QUIGLEY, H.A., SOMMER, A. (1995). Automated perimetry detects visual field loss before manual Goldmann perimetry. *Ophthalmology*, 102(1), p.21-26.

KELTNER, J.L., JOHNSON, C.A., ANDERSON, D.R., LEVINE, R.A., FAN, J., CELLO, K.E., QUIGLEY, H.A., BUDENZ, D.L., PARRISH, R.K., KASS, M.A., GORDON, M.O. (2006). The association between glaucomatous visual fields and optic nerve head features in the Ocular Hypertension Treatment Study. *Ophthalmology*, 113, p.1603-1612.

KELTNER, J.L., JOHNSON, C.A., BALESTRERY, F.G. (1979). Suprathreshold static perimetry. Initial clinical trials with Fieldmaster automated perimeter. *Archives of Ophthalmology*, 97, p.260-272.

KELTNER, J.L., JOHNSON, C.A., CELLO, K.E., BANDERMANN, S.E., FAN, J., LEVINE, R.A., KASS, M.A., GORDON, M.O.; Ocular Hypertension Treatment Study Group. (2007). Ocular Hypertension Study Group. Visual field quality control in the Ocular Hypertension Treatment Study (OHTS). *Journal of Glaucoma*, 16(8), p.665-669.

KELTNER, J.L., JOHNSON, C.A., QUIGG, J.M., CELLO, K.E., KASS, M.A., GORDON, M.O. (2000). Confirmation of visual field abnormalities in the Ocular Hypertension Treatment Study. Ocular Hypertension Treatment Study Group. *Archives of Ophthalmology*, 118, p.1187-1194.

KHAIRALLAH M, KAHLOUN R, BOURNE R, LIMBURG H, FLAXMAN SR, JONAS JB, KEEFFE J, LEASHER J, NAIDOO K, PESUDOVS K, PRICE H, WHITE RA, WONG TY, RESNIKOFF S, TAYLOR HR; Vision Loss Expert Group of the Global Burden of Disease Study. (2015). Number of People Blind or Visually Impaired by Cataract Worldwide and in World Regions, 1990 to 2010. *Investigative Ophthalmology and Visual Science*, 56(11), p.6762-6769.

KHUU, S. and KALLONIATIS, M. (2015a). Spatial summation across the visual field: implications for visual field testing. *Journal of Vision*, 15(1), p.1-15.

KHUU, S.K. and KALLONIATIS, M. (2015). Standard automated perimetry: determining spatial summation and its effect on contrast sensitivity across the visual field*. Investigative Ophthalmology and Visual Science*, 56, p.3565–3576.

KIM, J., DALLY, L.G., EDERER, F., GAASTERLAND, D.E., VANVELDHUISEN, P.C., BLACKWELL, B., SULLIVAN, E.K., PRUM, B., SHAFRANOV, G., BECK, A., SPAETH, G.L.; AGIS INVESTIGATORS (2004) The Advanced Glaucoma Intervention Study (AGIS): 14. Distinguishing progression of glaucoma from visual field fluctuations. *Ophthalmology*, 111, 2109-2116

KIM, Y.Y., KIM, J.S., SHIN, D.H., KIM, C., JUNG, H.R. (2001). Effect of cataract extraction on blue-on-yellow visual field. *American Journal of Ophthalmology*, 132(2), p. 217-220.

KING, A.J.W., TAGURI, A., WADOOD, A.C., AZUARA-BLANCO, A. (2002). Comparison of two fast strategies, SITA Fast and TOP, for the assessment of visual fields in glaucoma patients. *Graefe's Archive Clinical Experimental Ophthalmology*, 240, p.481–487.

KING, D., DRANCE, S.M., DOUGLAS, G., SCHULZER, M., WIJSMAN, K. (1986). Comparison of visual field defects in normal-tension glaucoma and high tension glaucoma. *American Journal of Ophthalmology*, 101, p.204-207.

KING-SMITH, P.E., GRIGSBY, S.S., VINGRYS, A.J., BENES, S.C., SUPOWIT, A. (1994). Efficient and unbiased modifications of the QUEST threshold method: theory, simulations, experimental evaluation and practical implementation. *Vision Research*, 34, p.885-912.

KITAZAWA, Y**.** and YAMAMOTO, T. (1997). Glaucomatous visual field defects: their characteristics and how to detect them. *Clinical Neuroscience (New York, N.Y.)***,** 4(5), p.279-283.

KLEIN, B.E., KLEIN, R., JENSEN, S.C. (1996). Visual sensitivity and age-related eye diseases: the Beaver Dam Eye Study. *Ophthalmic Epidemiology*, 3(1), p.47-55.

KLEIN, B.E., KLEIN, R., LINTON, K.L. (1992). Prevalence of age-related lens opacities in a population. The Beaver Dam Eye Study. *Ophthalmology*, 99(4), p.546-552.

KOCABEYOGLU, S., UZUN, S., MOCAN, M.C., BOZKURT, B., IRKEC, M., ORHAN, M. (2013). Comparison of visual field test results obtained through Humphrey matrix frequency doubling technology perimetry versus standard automated perimetry in healthy children. *Indian Journal of Ophthalmology*, 61, p. 576 – 579.

KOOK, M.S., YANG, S.J., KIM, S., CHUNG, J., KIM, S.T. and TCHAH, H. (2004). Effect of cataract extraction on Frequency Doubling Technology perimetry. *American Journal of Ophthalmology*, 138, p.85-90.

KOUCHEKI B, NOURI-MAHDAVI K, PATEL G, GAASTERLAND D, and CAPRIOLI, J. (2004). Visual field changes after cataract extraction: the AGIS experience. *American Journal of Ophthalmology*, 138(6), p.1022-1028.

KRAKAU, C.E. (1989). Visual field testing with reduced sets of test points. A computerized analysis. *Documenta Ophthalmologica*, 73, p.71-80.

KULZE, J., STEWART, W., SUTHERLAND, S. (1990). Factors associated with a learning effect in glaucoma patients using automated perimetry. *Acta Ophthalmologica*, 68, p.681-686.

KUTZKO, K.E., BRITO, C.F., WALL, M. (2000). Effect of instructions on conventional automated perimetry*. Investigative Ophthalmology and Visual Science,* 41, p.2006-2013*.*

LACHENMAYR, B.J., DRANCE, S.M., AIRAKSINEN, P.J. (1992). Diffuse field loss and diffuse retinal nerve-fiber loss in glaucoma. *German Journal of Ophthalmology*, 1(1), p.22-25.

LACHENMAYR, B.J., DRANCE, S.M., CHAUHAN, B.C., HOUSE, P.H., LALANI, S. (1991). Diffuse and localized glaucomatous field loss in light-sense, flicker and resolution perimetry. *Graefe's Archive for Clinical and Experimental Ophthalmology*, 229(3), p.267-273.

LACHKAR, Y., BARRAULT, O., LEFRANCOIS, A., DEMAILLY, P. (1998). Rapid tendency oriented perimeter (TOP) with the Octopus visual field analyzer. *Journal francais d'ophtalmologie*, 21, p.180-184.

LAM, B.L., ALWARD, W.L., KOLDER, H.E. (1991). Effect of cataract on automated perimetry. *Ophthalmology*, 98 (7), p.1066-1070.

LAMPARTER, J., ALIYEVA, S., SCHULZE, A., BERRES, M., PFEIFFER, N., HOFFMANN, E.M. (2013). Standard automated perimetry versus Matrix frequency doubling technology perimetry in subjects with ocular hypertension and healthy control subjects. *PLoS ONE*, 8(2): e57663.

LAMPARTER, J., SCHULZE, A., SCHUFF, A.C., BERRES, M., PFEIFFER, N., HOFFMANN, E.M. (2011). Learning curve and fatigue effect of flicker defined form perimetry. *American Journal of Ophthalmology*, 151(6), p.1057-1064.e1.

LANDERS, J., SHARMA, A., GOLDBERG, I., GRAHAM, S.(2003).A comparison of perimetric results with the Medmont and Humphrey perimeters.*British Journal of Ophthalmology*, 87(6), p.690-694.

LANDERS, J., SHARMA, A., GOLDBERG, I., GRAHAM, S.(2007). A comparison of global indices between the Medmont Automated Perimeter and the Humphrey Field Analyzer. *British Journal of Ophthalmology*, 91(10), p.1285-1287.

LANDERS, J., SHARMA, A., GOLDBERG, I., GRAHAM, S.L. (2010). Comparison of visual field sensitivities between the Medmont automated perimeter and theHumphrey field analyser. *Clinical and Experimental Ophthalmology*, 38(3), p.273–276.

LANGERHORST, C.T., VAN DEN BERG, T.J., GREVE, E.L. (1989). Is there general reduction of sensitivity in glaucoma? *International Ophthalmology*, 13(1-2), p.31-35.

LANGERHORST, C.T., VAN DEN BERG, T.J.T.P., VELDMAN, E., GREVE, E.L. (1987). Population study of global and local fatigue with prolonged threshold testing in automated perimetry. In: GREVE, E.L. and HEIJL, A.(eds.) *Seventh International Visual Field Symposium. DocumentaOphthalmologica Proceedings Series 49*. Dordrecht: MartinusNijhoff/Dr W Junk, 657 – 662.

LANSINGH, V.C., CARTER, M.J., MARTENS, M. (2007). Global cost-effectiveness of cataract surgery. *Ophthalmology*, 114, p.1670-1678.

LASCARATOS, J. and MARKETOS, S. (1988). A historical outline of Greek ophthalmology from the Hellenistic period up to the establishment of the first universities. *Documenta Ophthalmologica*, 68, p. 157– 169.

LAU, L.I., LIU, C.J., CHOU, J.C., HSU, W.M., LIU, J.H.K. (2003). Patterns of visual field defects in chronic angle-closure glaucoma with different disease severity. *Ophthalmology,* 110, p.1890–1894

LE, P.V., ZHANG, X., FRANCIS, B.A., VARMA, R., GREENFIELD, D.S., SCHUMAN, J.S., LOEWEN, N., HUANG, D.; Advanced Imaging for Glaucoma Study Group. (2015). Advanced imaging for glaucoma study: design, baseline characteristics, and inter-site comparison. *American Journal of Ophthalmology,*159(2), p. 393 – 403.e2.

LEBLANC, E.P. and BECKER, B. (1971). Peripheral nasal field defects. *American Journal of Ophthalmology*, 72(2), p.415-419.

LEE, A.J., WANG, J.J., ROCHTCHINA, E., HEALEY, P., CHIA, E.M., MITCHEL, P. (2003). Patterns of glaucomatous visual field defects in an older population: the Blue Mountains Eye Study. *Clinical and Experimental Ophthalmology*, 31(4), p.331-335.

LEEPRECHANON, N., GIANGIACOMO, A., FONTANA, H., HOFFMAN, D., CAPRIOLI, J. (2007). Frequency-doubling perimetry: comparison with standard automated perimetry to detect glaucoma. *American Journal of Ophthalmology*, 143, p.263-271.

LEFFLER, C.T., SCHWARTZ, S.G., HADI, T.M., SALMAN, A., VASUKI, V. (2015). The early history of glaucoma: the glaucous eye (800 BC to 1050 AD). *Clinical Ophthalmology*, 9, p.207-215.

LEITE,M.T., ZANGWILL, L.M., WEINREB, R.N., RAO, H.L., ALENCAR, L.M., MEDEIROS, F.A. (2012). Structure-function relationships using the Cirrus spectral domain optical coherence tomograph and standard automated perimetry. *Journal of Glaucoma*, 21(1), p.49-54.

LELKENS, A.M. AND ZUIDEMA, P. (1983). Increment thresholds with various low background intensities at different locations in the peripheral retina. *Journal of the Optical Society of America*, 73, p.1372-1378.

LESKE, M.C., CONNELL, A.M., SCHACHAT, A.P., HYMAN, L. (1994) The Barbados Eye Study. Prevalence of open-angle glaucoma. *Archives of Ophthalmology,* 112, p.821-829.

LESKE, M.C., HEIJL, A., HYMAN, L., BENGTSSON, B. (1999). Early Manifest Glaucoma Trial: design and baseline data. *Ophthalmology,* 106, p.2144-2153.

LEWIS, R.A., HAYREH, S.S., PHELPS, C.D. (1983). Optic disk and visual field correlations in primary open-angle and low-tension glaucoma. *American Journal of Ophthalmology*, 96, p.148-152.

LEWIS, R.A., JOHNSON, C.A., KELTNER, J.L., LABERMEIER, P.K. (1986). Variability of quantitative automated perimetry in normal observers. *Ophthalmology*, 93, p.878-881.

LI, E.Y., LIU, Y., ZHAN, X., LIANG, Y.B., ZHANG, X., ZHENG, C., JHANJI, V., XU, P., CHANG, D.F., LAM, D.S. (2013). Prevalence of blindness and outcomes of cataract surgery in Hainan Province in South China. *Ophthalmology,* 120(11), p.2176-2183.

LICHTER, P.R., MUSCH, D.C., GILLESPIE, B.W., GUIRE, K.E., JANZ, N.K., WREN, P.A., MILLS, R.P. (2001). Interim clinical outcomes in the Collaborative Initial Glaucoma Treatment Study comparing initial treatment randomized to medications or surgery. *Ophthalmology*, 108, p.1943-1953.

LIN, A., HOFFMAN, D., GAASTERLAND, D.E., CAPRIOLI, J. (2003) Neural networks to identify glaucomatous visual field progression. *American Journal of Ophthalmology*, 135, p.49-54.

LISBOA, R., LEITE, M.T., ZANGWILL, L.M., TAFRESHI, A., WEINREB, R.N., MEDEIROS, F.A.. (2012). Diagnosing preperimetric glaucoma with spectral domain optical coherence tomography. *Ophthalmology*, 119(11), p.2261-2269.

LIU, S., LAM, S., WEINREB, R.N., YE, C., CHEUNG, C.Y., LAI, G., LAM, D.S., LEUNG, C.K. (2011). Comparison of standard automated perimetry, frequency-doubling technology perimetry, and short-wavelength automated perimetry for detection for glaucoma. *Investigative Ophthalmology and Visual Science*, 52, p.7325-7331.

LOPEZ-PENA, M.J., FERRERAS, A., LARROSA, J.M., POLO, V., PABLO, L.E. (2011). Relationship between standard automated perimetry and retinal nerve fiber layer parameters obtained with optical coherence tomography. *Journal of Glaucoma*, 20, p.422-432

LORCH, L., DIETRICH, T.J., Schwabe, R., SCHIEFER, U. (2001). Comparison of local differential luminance sensitivity (dls) between Oculus Twinfield Perimeter and Humphrey Field Analyzer 630 (HFA I) in normal volunteers of varying ages (abstract). *Klinische Monatsblätter für Augenheilkunde*, 218(12), p.782-794.

LOU, D., YIP, K.W., YAO, P., CHEN, Y., CHEN, Z., ZHOU, X., (2011). The study of eyestrain and lighting conditions. IN: Proceeding of 27th Session of the CIE, Sun City, South Africa, p.52–58

LUITHARDT, A.F., MEISNER, C., MONHART, M., KRAPP, E., MAST, A., SCHIEFER, U. (2015). Validation of a new static perimetric thresholding strategy (GATE). *British Journal of Ophthalmology*, 99, p.11-15.

LUNDSTROM, M., GOH, P.P., HENRY. Y., SALOWI, M.A., BARRY, P., MANNING, S., ROSEN, P., STENEVI. U. (2015). The changing pattern of cataract surgery indications: a 5-year study of 2 cataract surgery databases. *Ophthalmology*, 122, p.31-38.

MADDESS, T. (2011). The influence of sampling errors on test-retest variability in perimetry. *Investigative Ophthalmology and Visual Science*, 52, p.1014-1022.

MAEDA, H., NAKAURA, M. and NEGI, A. (2000). New perimetric threshold test algorithm with dynamic strategy and tendency oriented perimetry (TOP) in glaucomatous eyes. *Eye*, 5, p.747-751.

MAGNUS, H. (1998). Ophthalmology of the ancients (Vol. 4, part 1) (R.L. Waugh, Trans.). Belgium: Wayenborgh.

MARDIN, C.Y., HORN, F.K., JONAS, J.B., BUDDE, W.M. (1999). Preperimetric glaucoma diagnosis by confocal scanning laser tomography of the optic disc. *British Journal of Ophthalmology*, 83(3), p.299-304.

MARDIN, C.Y., PETERS, A., HORN, F., JÜNEMANN, A.G., Lausen, B.(2006). Improving glaucoma diagnosis by the combination of perimetry and HRT measurements. *Journal of Glaucoma*, 15, p.299-305.

MARRA, G. and FLAMMER, J. (1991). The learning and fatigue effect in automated perimetry. *Graefe's Archive for Clinical and Experimental Ophthalmology*, 229, p.501-504.

MASLIN, J.S., MANSOURI, K., DORAIRAJ, S.K. (2015). HRT for the Diagnosis and Detection of Glaucoma Progression. *The Open Ophthalmology Journal*, 9, p.58-67

McCARTY, C.A., MUKESH, B.N., FU, C.L., TAYLOR, H.R. (1999). The epidemiology of cataract in Australia. *American Journal of Ophthalmology*, 128(4), p.446-465.

MEDEIROS, F.A., SAMPLE, P.A., ZANGWILL, L.M., LIEBMANN, J.M., GIRKIN, C.A., WEINREB, R.N. (2006). A statistical approach to the evaluation of covariate effects on the receiver operating characteristic curves of diagnostic tests in glaucoma. *Investigative Ophthalmology and Visual Science*, 47, p.2520-2527.

MEYER, D.R., STERN, J.H., JARVIS, J.M, LININGER, L.L. (1993). Evaluating the visual field effects of blepharoptosis using automated static perimetry. *Ophthalmology*, 100, p.651-659.

MIGLIOR S, ZEYEN T, PFEIFFER N, CUNHA-VAZ J, TORRI V, ADAMSONS I. (2002). The European glaucoma prevention study design and baseline description of the participants. *Ophthalmology*, 109, p.1612-1621.

MIGLIOR S, ZEYEN T, PFEIFFER N, CUNHA-VAZ J, TORRI V, ADAMSONS, I. (2005). Results of the European Glaucoma Prevention Study. *Ophthalmology,* 112, p.366-375.

MIGLIOR, S., PFEIFFER, N., TORRI, V., ZEYEN, T., CUNHA-VAZ, J., ADAMSONS I. (2007). Predictive factors for open-angle glaucoma among patients with ocular hypertension in the European Glaucoma Prevention Study. *Ophthalmology*, 114(1), p.3-9.

MILLER, K.N., SHIELDS, M.B., OLLIE, A.R. (1989). Automated kinetic perimetry with two peripheral isopters in glaucoma. *Archives of Ophthalmology*, 107(9), p.1316-1320.

MILLS, R.P. (1984). A comparison of Goldmann, Fieldmaster 200, and Dicon AP2000 perimeters used in a screening mode. *Ophthalmology*, 91, p.347-354.

MITCHELL, P., SMITH, W., ATTEBO, K., HEALEY, P.R. (1996). Prevalence of open-angle glaucoma in Australia: the Blue Mountains Eye Study. *Ophthalmology*, 103(10), p.1661-1669.

MOKHLES, P., SCHOUTEN, J.S., BECKERS, H.J., AZUARA-BLANCO, A., TUULONEN, A., WEBERS, C.A. (2016). A systematic review of end-of-life visual impairment in open-angle glaucoma: an epidemiological autopsy. *Journal of Glaucoma*, 25(7), p.623-628.

MORALES, J., WEITZMAN, M., GONZALEZ DE LA ROSA, M. (2000). Comparison between Tendency-Oriented Perimetry (TOP) and Octopus threshold perimetry. *Ophthalmology*, 107, p.134–142.

MOSS, I.D. and WILD, J.M. (1994). The influence of induced forward light scatter on the normal blue-on-yellow perimetric profile. *Graefe's Archive for Clinical and Experimental Ophthalmology*, 232, p.409-414.

MOSS, I.D., WILD, J.M., WHITAKER, D.J. (1995). The influence of age-related cataract on blue-on-yellow perimetry. *Investigative Ophthalmology and Visual Science*, 36, p.764-773.

MOTOLKO, M., DRANCE, S.M., DOUGLAS, G.R. (1982). Visual field defects in low-tension glaucoma. Comparison of defects in low-tension glaucoma and chronic open angle glaucoma. *Archives of Ophthalmology*, 100(7), p.1074-1077.

MUELLER, C.G. (1951). Frequency of seeing functions for intensity discrimination of various levels of adapting intensity. *The Journal of General Physiology*,34, p.463–474.

MULHOLLAND, P.J., REDMOND, T., GARWAY-HEATH, D.F., ZLATKOVA, M.B., ANDERSON, R.S.(2015). Spatiotemporal summation of perimetric stimuli in early glaucoma. *Investigative Ophthalmology and Visual Science*, 56, p.6473–6482.

MULHOLLAND, P.J., REDMOND, T., GARWAY-HEATH, D.F., ZLATKOVA, M.B., ANDERSON, R.S.(2015a). Estimating the critical duration for temporal summation of standard achromatic perimetric stimuli. *Investigative Ophthalmology and Visual Science*, 56, p.431-437.

MUSCH, D.C., GILLESPIE, B.W., NIZIOL, L.M., JANZ, N.K., WREN, P. A., ROCKWOOD, E.J., LICHTER, P.R. (2006). Cataract extraction in the Collaborative Initial Glaucoma Treatment Study: incidence, risk factors, and the effect of cataract progression and extraction on clinical and quality-of-life outcomes. *Archives of Ophthalmology*, 124, p.1694-1700.

MUSCH, D.C., LICHTER, P.R., GUIRE, K.E., STANDARDI, C.L. (1999). The Collaborative Initial Glaucoma Treatment Study: study design, methods, and baseline characteristics of enrolled patients. *Ophthalmology***,** 106(4), p.653-662.

NA, K.-S., PARK, Y.-G., HAN, K., MOK, J. W. and JOO, C.-K. (2014). Prevalence of and Risk Factors for Age-Related and Anterior Polar Cataracts in a Korean Population. *PLoS ONE*, *9*(6), e96461.

NELSON-QUIGG, J., TWELKER, J.D., JOHNSON, C.A. (1989). Response properties of normal observers and patients during automated perimetry. *Archives of Ophthalmology* 107(11), p.1612-1615.

NEWKIRK, M.R., GARDINER, S.K., DEMIREL, S., JOHNSON, C.A. (2006). Assessment of false positives with the Humphrey Field Analyzer II perimeter with the SITA Algorithm. *Investigative Ophthalmology & Visual Science*, 47, p.4632-4637.

NG, M., SAMPLE, P.A., PASCUAL, J.P., ZANGWILL, L.M., GIRKIN, C.A., LIEBMANN, J.M., WEINREB, R.N., RACETTE, L. (2012). Comparison of visual field severity classification systems for glaucoma. *Journal of Glaucoma*, 21(8), p.551-561.

NORDMANN, J.P., BRION, F., HAMARD, P., MOUTON-CHOPIN, D. (1998). Evaluation of the Humphrey perimetry programs SITA Standard and SITA Fast in normal probands and patients with glaucoma. *Journal françaisd'ophtalmologie*, 21(8), p.549-554.

NORDMANN, J.P., DENIS, P., NGUER, Y., MOUTON-CHOPIN, D., SARAUX, H. (1994). Static threshold visual field in glaucoma with the Fastpac algorithm of the Humphrey Field Analyser. Is the gain in examination time offset by any loss of information? *European Journal of Ophthalmology*, 4, p.105-110.

NOURI-MAHDAVI K, HOFFMAN D, RALLI M, CAPRIOLI J (2007) Comparison of methods to predict visual field progression in glaucoma. *Archives of Ophthalmology,* 125, p. 1176–1181.

NOURI-MAHDAVI, K. (2014). Selecting visual field tests and assessing visual field deterioration in glaucoma. *Canadian Journal of Ophthalmology*, 49, p. 497–505.

NOURI-MAHDAVI, K., CAPRIOLI, J., COLEMAN, A.L., HOFFMAN, D., GAASTERLAND, D. (2005). Pointwise linear regression for evaluation of visual field outcomes and comparison with advanced glaucoma intervention study methods. *Archives of Ophthalmology*, 123(2), p.193-199.

NOURI-MAHDAVI, K., HOFFMAN, D., COLEMAN, A.L., LIU, G., LI, G., GAASTERLAND, D., CAPRIOLI, J. (2004). Predictive factors for glaucomatous visual field progression in the Advanced Glaucoma Intervention Study. *Ophthalmology*, 111, p.1627-1635.

NOURI-MAHDAVI, K., NASSIRI, N., GIANGIACOMO, A., CAPRIOLI, J. (2011). Detection of visual field progression in glaucoma with standard achromatic perimetry: a review and practical implications. *Graefe's Archive Clinical Experimental Ophthalmology*, 249, p.1593-1616.

NOURI-MAHDAVI, K., SUPAWAVEJ, C., BITRIAN, E., GIACONI, J.A., LAW, S.K., COLEMAN, A.L., CAPRIOLI, J. (2011a). Patterns of damage in chronic angle-closure glaucoma compared to primary open-angle glaucoma. *American Journal of Ophthalmology*, 152, p.74-80.e2.

NOVAK-LAUS, K., KNEZEVIC, T., MATEJCIC, A., IVEKOVIC, R., ZORIC-GEBER, M., KORSIC, J., MANDIC, Z. (2007). Effect of cataract extraction on visual field in patients with open angle glaucoma. *Acta Clinica Croatica*, 46 (suppl 1), p.31-35.

NOWAK, M.S. and SMIGIELSKI, J. (2015). The prevalence of age-related eye diseases and cataract surgery among older adults in the city of Lodz, Poland. *Journal of Ophthalmology*, Article ID 605814, 7 pages

O'BRIEN, C., POINOOSAWMY, D., WU, J., HITCHINGS, R. (1994). Evaluation of the Humphrey FASTPAC threshold program in glaucoma. *British Journal of Ophthalmology*, 78, p.513-514.

ODDEN, J.L., MIHAILOVIC, A., BOLAND, M.V., FRIEDMAN, D.S., WEST, S.K., RAMULU, P.Y. (2016). Evaluation of central and peripheral visual field concordance in glaucoma. *Investigative Ophthalmology and Visual Science*, 57, p.2797-2804.

OLIVER, J.E., HATTENHAUER, M.G., HERMAN, D., HODGE, D.O., KENNEDY, R., FANG-YEN, M., 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, p.764-772.

OLSSON, J. and ROOTZEN, H. (1994): An image model for quantal response analysis in perimetry. *Scandinavian Journal of Statistics*, 21, p.375-387.

OLSSON, J., HEIJL, A., BENGTSSON, B., ROOTZEN, H. (1993). Frequency-of-seeing in computerised perimetry. IN: MILLS, R.P., ed. *Perimetry Update 1992/ 1993*. Amsterdam: Kugler; 551-556.

OLSSON, J., BENGTSSON, B., HEIJL, A., ROOTZEN, H. (1997). An improved method to estimate frequency of false positive answers in computerized perimetry. *Acta Ophthalmologica Scandinavica*, 75(2), p.181-183.

OTAROLA, F., CHEN, A., MORALES, E., YU, F., AFIFI, A., CAPRIOLI, J. (2016). Course of glaucomatous visual field loss across the entire perimetric range. *JAMA Ophthalmology*, 134(5), p.496-502.

PAPP, A., KIS, K., NÉMETH, J. (2001). Conversion Formulas between Automated-Perimetry Indexes as Measured by Two Different Types of Instrument. *Ophthalmologica*, 215, p.87-90

PARK, J. Y., HA, R. Y., RYU, V., KIM, E., JUNG, Y. C. (2013). Effects of color temperature and brightness on electroencephalogram alpha activity in a polychromatic light-emitting diode. *Clinical Psychopharmacology and Neuroscience*, 11(3), p.126-131.

PASCOLINI, D. and MARIOTTI, S.P. (2012). Global estimates of visual impairment: 2010. *British Journal of Ophthalmology*, 96, p.614-618.

PASCOLINI, D., MARIOTTI, S.P., POKHAREL, G.P., PARARAJASEGARAM, R., ETYA'ALE, D., NÉGREL, A.-D., RESNIKOFF, S. (2004). 2002 global update of available data on visual impairment: a compilation of population-based prevalence studies. *Ophthalmology Epidemiology*, 11(2), p.67-115.

PATEL, A., WOLLSTEIN, G., ISHIKAWA, H., SCHUMAN, J.S. (2007). Comparison of visual field defects using Matrix perimetry and standard achromatic perimetry. *Ophthalmology*, 114, p.480-487.

PEARSON, P.A., BALDWING, L.B., SMITH, T.J. (1990). The relationship of mean defect to corrected loss variance in glaucoma and ocular hypertension. *Ophthalmologica,* 200, p.16–21.

PENA-BETANCOR, C., GONZALEZ-HERNANDEZ, M., FUMERO-BATISTA,F., SIGUT, J., MEDINA-MESA, E., ALAYON, S., GONZALEZ DE LA ROSA, M. (2015). Estimation of the relative amount of haemoglobin in the cup and neuroretinal rim using stereoscopic color fundus images. *Investigative Ophthalmology and Visual Science*, 56 (3), p.1562-1568.

PESUDOVS, K. and ELLIOTT, D.B. (2003). Refractive error changes in cortical, nuclear, and posterior subcapsular cataracts. *British Journal of Ophthalmology*, 87, p.964-967.

PETRACO, R., DEHBI, H., HOWARD, J.P., SHUN-SHIN, M.J., SEN, S., NIJJER, S.S., MAYET, J., DAVIES, J.E., FRANCIS, D.P. (2018). Effects of disease severity distribution on the performance of quantitative diagnostic methods and proposal of a novel 'V-plot' methodology to display accuracy values. *Open Heart*, 5, p.e000663.

PETERS, D., BENGTSSON, B., HEIJL, A. (2013). Lifetime risk of blindness in open-angle glaucoma. *American Journal of Ophthalmology*, 156(4), p.724-730.

PHU, J., KHUU, S.K., YAPP, M., ASSAAD, N., HENNESSY, M.P., KALLONIATIS, M. (2017). The value of visual field testing in the era of advanced imaging: clinical and psychophysical perspectives. *Clinical and Experimental Optometry, 100*, p. 313-332.

PIERRE-FILHO, P.T., SCHIMITI, R.B., DE VASCONCELLOS, J.P., 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, p.345-350.

POLO, V., LARROSA, J.M., PINILLA, I., GONZALVO, F., FERRERAS, A., HONRUBIA, F.M. (2002). Glaucomatous damage patterns by short-wavelength automated perimetry (SWAP) in glaucoma suspect. *European Journal of Ophthalmology*, 12(1), p.49-54.

POSNER, A. and SCHLOSSMAN, A. (1948). Development of changes in visual fields associated with glaucoma. *Archives of Ophthalmology*, 39, p.623-639.

PRAJAPATI, B., DUNNE, M., ARMSTRONG, R. (2010). Sample size estimation and statistical power analyses. *Optometry Today*, 16(7), p.10-18

PYE, D., HERSE, P., NGUYEN, H. (1999). Conversion factor for comparison of data from Humphrey and Medmont automated perimeters. *Clinical and Experimental Optometry*, 82, p.11-14.

QUIGLEY, H.A. (1996). Number of people with glaucoma worldwide. *British Journal of Ophthalmology*, 80, p.389-393.

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

RACETTE, L., FISCHER, M., BEBIE, H., HOLLÓ, G., JOHNSON, C.A., MATSUMOTO, C. (2016). *Visual field digest: a guide to perimetry and the Octopus perimeter.* 6th edition. Köniz, Switzerland: Haag-Streit AG.

RACETTE, L., MEDEIROS, F.A., ZANGWILL, L.M., NG, D., WEINREB, R.N., SAMPLE, P.A. (2008). Diagnostic accuracy of the Matrix 24-2 and original N-30 frequency-doubling technology tests compared with standard automated perimetry. *Investigative Ophthalmology and Visual Science*, 49(3), p. 954-960.

RAO, G.N., KHANNA, R., PAYAL, A. (2011). The global burden of cataract. *Current Opinion in Ophthalmology*, 22, p.4-9.

RAO, H.L., JONNADULA, G.B., ADDEPALLI, U.K., SENTHIL, S., GARUDADRI, C.S. (2013). Effect of cataract extraction on visual field index in glaucoma. *Journal of Glaucoma*, 22(2), p.164-168.

RAO, H.L., YADAV, R.K., BEGUM, V.U., ADDEPALLI, U.K., CHOUDHARI, N.S., SENTHIL, S., GARUDADRI, C.S. (2015). Role of visual field reliability indices in ruling out glaucoma. *JAMA Ophthalmology*, 133(1), p.40-44.

RAO, H.L., RAVEENDRAN, S., JAMES, V., DASARI, S., PALAKURTHY, M., REDDY, H.B., PRADHAN, Z.S., RAO, D.A., PUTTAIAH, N.K., DEVI, S. (2017). Comparing the performance of Compass perimetry with Humphrey Field Analyzer in eyes with glaucoma. *Journal of Glaucoma,* 26(3), p.292-297

REALINI, T. (2011). A history of glaucoma pharmacology. *Optometry and Vision Science*, 88, p.36-38.

REDDY, G.R. (2010). *Practical Guide to Interpret Visual Fields*. 3rd Edition. New Dehli: Jaypee Brothers

REDMOND, T., GARWAY-HEATH, D.F., ZLATKOVA, M.B., ANDERSON, R.S. (2010). Sensitivity loss in early glaucoma can be mapped to an enlargement of the area of complete spatial summation. *Investigative Ophthalmology and Visual Science*, 51(12), p.6540-6548.

REDMOND, T., O'LEARY, N., HUTCHISON, D.M., NICOLELA, M.T., ARTES, P.H., CHAUHAN, B.C. (2013a). Visual field progression with frequency-doubling matrix perimetry and standard automated perimetry in patients with glaucoma and in healthy controls. *JAMA Ophthalmology*, 131, p.1565-1572.

REDMOND, T., ZLATKOVA, M.B., GARWAY-HEATH, D.F., ANDERSON, R.S. (2010a). The effect of age on the area of complete spatial summation for chromatic and achromatic stimuli. *Investigative Ophthalmology and Visual Science*, 51, p.6533–6539.

REDMOND, T., ZLATKOVA, M.B., VASSILEV, A., GARWAY-HEATH, D.E., ANDERSON, R.S. (2013). Changes in Ricco's area with background luminance in the S-cone pathway. *Optometry and Vision Science*, 90, p.66-74.

REHMAN SIDDIQUI, M.A., KHAIRY, H.A. and AZUARA-BLANCO, A. (2007). Effect of cataract extraction on SITA perimetry in patients with glaucoma. *Journal of Glaucoma*, 16, p.205-208.

REMKY, A. and AREND, O. (2000). Clinical experiences with the "Swedish interactive threshold algorithm (SITA)". *Klinische Monatsblätterfür Augenheilkunde*, 216 (3), p.143-147. [Article in German]

RESNIKOFF, S., PASCOLINI, D., ETYA'ALE, D., KOCUR, I., PARARAJASEGARAM, R., POKHAREL, G.P., MARIOTTI, S.P. (2004). Global data on visual impairment in the year 2002. *Bull. World Health Organ*, 82, p.844-851.

RHEE, K., KIM, Y.Y., NAM, D.H., JUNG, H.R. (2001). Comparison of visual field defects between primary open-angle glaucoma and chronic primary angle-closure glaucoma in the early or moderate stage of the disease. *Korean Journal of Ophthalmology*, 15, p.27-31.

RICCO, A. (1877). Relazione fra il minimo angolo visuale e l'intensita` luminosa. Memorie della Regia Academia di Scienze, lettere ed arti in Modena, 17, p.47Y160.

RISSE, J.F., DUMAUSE, A., HUE, B. (1999). [Comparative study of 2 classifications of glaucomatous perimetric deficits]. *Journal français d'ophtalmologie*, 22(7), p.738-742.

ROGGEN, X., HERMAN, K., VAN MALDEREN, L., DEVOS, M., SPILEERS, W. (2001). Different strategies for Humphrey automated perimetry: FASTPAC, SITA standard and SITA fast in normal subjects and glaucoma patients. *Bulletin de la Societe Belged'Ophtalmologie*, 279, p.23-33.

ROSSETTI, L., FOGAGNOLO, P., MIGLIOR, S., CENTOFANTI, M., VETRUGNO, M., ORZALESI, N. (2006). Learning effect of short-wavelength automated perimetry in patients with ocular hypertension. *Journal of Glaucoma*, 15(5), p.399-404

ROUNTREE, L., MULHOLLAND, P.J., ANDERSON, R.S., GARWAY-HEATH, D.F., MORGAN, J.E., REDMOND, T. (2018). Optimising the glaucoma signal or noise ratio by mapping changes in spatial summation with area-modulated perimetric stimuli. *Scientific Reports*, 8, p.2172.

ROWE, F.J., WISHART, M., SPENCER, S. (2014). Perimetry comparisons for Octopus G Top and Dynamic programmes versus Humphrey 24-2 SITA Fast and SITA Standard programmes *Ophthalmology Research: An International Journal*, 2(1), p.24-42

RUDNICKA, A.R.and EDGAR, D.F. (1996). Automated static perimetry in myopes with peripapillary crescents--Part II. *Ophthalmic and Physiological Optics*, 16(5), p.416-429.

SAKATA, L.M., DELEON-ORTEGA, J., ARTHUR, S.N., MONHEIT, B.E., GIRKIN, C.A. (2007). Detecting visual function abnormalities using the Swedish interactive threshold algorithm and matrix perimetry in eyes with glaucomatous appearance of the optic disc. *Archives of Ophthalmology*, 125, p.340-345.

SALMON, J.F., MERMOUD, A., IVEY, A., SWANEVELDER, S.A., HOFFMAN, M. (1993). The prevalence of primary angle closure glaucoma and open angle glaucoma in Mamre, Western Cape, South Africa. *Archives of Ophthalmology*, 111, p.1263-1269.

SAMPLE PA, CHAN K, BODEN C, LEE, T., BLUMENTHAL, E.Z., WEINREB, R.N., BERND, A., PASCUAL, J., HAO, J., SEJNOWSKI, T., GOLDBAUM, M.H. (2004). Using unsupervised learning with variational Bayesian mixture of factor analysis to identify patterns of glaucomatous visual field defects. *Investigative Ophthalmology and Visual Science,* 45, p.2596-2605.

SAMPLE, P.A. (2001). What does functional testing tell us about optic nerve damage? *Survey of Ophthalmology*, 45(Suppl3), p.S319-S324

SAMPLE, P.A., BOSWORTH, C.F., WEINREB, R.N. (2000). The loss of visual function in glaucoma. *Seminars in Ophthalmology*, 15, p.182-193.

SAMPLE, P.A., DANNHEIM, F., ARTES, P.H., DIETZSCH, J., HENSON, D., JOHNSON, C.A., NG, M., SCHIEFER, U., WALL, M.; the IPS Standards Group (2011). Imaging and Perimetry Society standards and guidelines. *Optometry and Vision Science*, 88, p.4–7.

SAMPLE, P.A., MEDEIROS, F.A., RACETTE, L., PASCUAL, J.P., BODEN, C., ZANGWILL, L.M., BOWD, C., WEINREB, R.N. (2006). Identifying glaucomatous vision loss with visual-function-specific perimetry in the Diagnostic Innovations in Glaucoma Study. *Investigative Ophthalmology and Visual Science*, 47(8), p.3381-3389.

SANABRIA, O., FEUER, W.J., ANDERSON, D.R. (1991). Pseudo-loss of fixation in automated perimetry. *Ophthalmology*, 98, p.76–78.

SAUNDERS, R.M. (1975). The critical duration of temporal summation in the human central fovea. *Vision Research*, 15, p.699–703.

SCHAUMBERGER, M., SCHAFER, B., LACHENMAYR, B.J. (1995). Glaucomatous visual fields. Fastpac versus Full Threshold strategy of the Humphrey Field Analyzer. *Investigative Ophthalmology and Visual Science*, 36, p.1390-1397.

SCHIEFER, U., PASCUAL, J.P., EDMUNDS, B., FEUDNER, E., HOFFMANN, E.M., JOHNSON, C.A., LAGRÈZE, W.A., PFEIFFER, N., SAMPLE, P.A., STAUBACH, F., WELEBER, R.G., VONTHEIN, R., KRAPP, E., PAETZOLD, J. (2009). Comparison of the new perimetric GATE strategy with conventional full-threshold and SITA Standard strategies. *Investigative Ophthalmology and Visual Science*, 50, p.488-494.

SCHIEFER, U., PÄTZOLD, J., DANNHEIM, F. (2005). Konventionalle Perimetrie. Teil I: Einführung – Grundbegriffe. *Der Ophthalmologe*, 102(6), p. 627-646.

SCHIMITI, R.B., AVELINO, R.R., KARA-JOSE, N., COSTA, V.P. (2002). Full-threshold versus Swedish Interactive Threshold Algorithm (SITA) in normal individuals undergoing automated perimetry for the first time. *Ophthalmology*, 109, p.2084-2092.

SCHULZER, M., MIKELBERG, F.S., DRANCE, S.M. (1987). A study of the value of the central and peripheral isoptres in assessing visual field progression in the presence of paracentral scotomas measurements. *British Journal of Ophthalmology*, 71(6), p.422-427.

SEAMONE, C., LEBLANC, R., RUBILLOWICZ, M., MANN, C., ORR, A.. (1988). The value of indices in the central and peripheral visual fields for the detection of glaucoma. *American Journal of Ophthalmology*, 106(2), p.180-185.

SEARLE, A.E.T., WILD, J.M., SHAW, D.E., O'NEILL, E.C. (1991). Time-related variation in normal automated static perimetry. *Ophthalmology,* 98, p.701-707.

SEKHAR, G. C., NADUVILATH, T. J., LAKKAI, M., JAYAKUMAR, A. J., PANDI, G. T., MANDAL, A. K., & HONAVAR, S. G. (2000). Sensitivity of Swedish interactive threshold algorithm compared with standard full threshold algorithm in Humphrey visual field testing. *Ophthalmology*, 107, p.1303-1308.

SHAFI, A., SWANSON, W.H., DUL, M.W. (2011). Structure and function in patients with glaucomatous defects near fixation. *Optometry Vision Science*, 88(1), p.130-139.

SHAH, N.N., BOWD, C., MEDEIROS, F.A., WEINREB, R.N., SAMPLE, P.A., HOFFMANN, E.M., ZANGWILL, L.M.. (2006). Combining structural and functional testing for detection of glaucoma. *Ophthalmology*, 113, p.1593-1602.

SHANDIZ, J.H.; DERAKHSHAN, A., DANESHYAR, A., AZIMI, A., MOGHADDAM, H.O., YEKTA, A.A., YAZDI, S.H.H., ESMAILY, H. (2011). Effect of Cataract Type and Severity on Visual Acuity and Contrast Sensitivity. *Journal of Ophthalmic and Vision Research,* 6 (1), p. 26-31.

SHARMA, A.K., GOLDBERG, I., GRAHAM, S.L., MOHSIN, M. (2000). Comparison of the Humphrey Swedish Interactive Thresholding Algorithm (SITA) and full threshold strategies. *Journal of Glaucoma*, 9(1), p.20-27.

SHARMA, P., SAMPLE P.A., ZANGWILL, L.M., SCHUMAN, J.S. (2008). Diagnostic tools for glaucoma detection and management. *Survey of Ophthalmology*, 53, p.S17-S32.

SHELAT, B.B. and RAO, B.S. (2009). Automated perimetry in glaucoma. IN GUPTA, A.K. and KRISHNA, V. (eds). *Clinical Ophthalmology: Contemporary Perspective 9th ed*. Dehli: Elsevier, p.335-370.

SHERAFAT, H., SPRY, P.G.D., WALDOCK, A., SPARROW, J.M., DIAMOND, J.P. (2003). Effect of a patient training video on visual field test reliability. *British Journal of Ophthalmology*, 87, p.153-156

SHIRATO, S., INOUE, R., FUKUSHIMA, K., SUZUKI, Y. (1999). Clinical evaluation of SITA: a new family of perimetric testing strategies. *Graefe's Archive for Clinical and Experimental Ophthalmology*, 237(1), p.29-34.

SIHOTA, R., GUPTA, V., TULI, D., SHARMA, A., SONY, P., SRINIVASAN, G. (2007). Classifying patterns of localized glaucomatous visual field defects on automated perimetry. *Journal of Glaucoma,* 16(1), p.146-152.

SLOAN, L.L. (1961). Area and luminance of test object as variables in examination of the visual field by projection perimetry. *Vision Research*, 1, p.121-138, IN1 – IN2.

SLOAN, L.L. and BROWN, D.J. (1962). Area and luminance of test object as variables in projection perimetry: clinical studies of photometric dysharmony. *Vision Research*, 2, p.527-541.

SMITH, G. (2002). Disability glare and its clinical significance. *Optometry Today*. Available from:
http://ot.kenthouse.com/uploads/articles/f5dc196f4748c05151e9ad4e9d2bf5f8_smith20020418.pdf [Assessed 24 Jan 2017]

SMITH, S.D., KATZ, J., QUIGLEY, H.A. (1997). Effect of cataract extraction on the results of automated perimetry in glaucoma. *Archives of Ophthalmology*, 115(12), p.1515-1519.

SOLIMAN, M.A.E., DE JONG, L.A.M S., ISMAEIL, A.A., VAN DEN BERG, T.J.T.P., DE SMET, M.D.(2002) Standard achromatic perimetry, short wavelength automated perimetry, and frequency doubling technology for detection of glaucoma damage. *Ophthalmology,*109, p.444-454

SOMMER, A., DUGGAN, C., AUER, C., ABBEY, H. (1985). Analytic approaches to the interpretation of automated threshold perimetric data for the diagnosis of early glaucoma. *Transactions of the American Ophthalmological Society*, 83, p.250-267.

SOMMER, A., ENGER, C., WITT, K. (1987). Screening for glaucomatous visual field loss with automated threshold perimetry. *American Journal of Ophthalmology*, 103(5), p.681-684.

SONG, P., WANG, H., THEODORATOU, E., CHAN, K.Y., RUDAN, I. (2018). The national and subnational prevalence of cataract and cataract blindness in China: a systematic review and meta-analysis. *Journal of Global Health*, 8(1): 010804

SPAHR, J. and FANKHAUSER, F. (1974). OCTOPUS – an automated perimeter. *Rev Sensory Disability*, 18, p.5–8.

SPRY, P.G. and JOHNSON. C.A. (2001). Senescent changes of the normal visual field: an age-old problem. *Optometry and Vision Science*, 78(6), p.436-441

SPRY, P.G.D. and HARPER, R.A. (2010). *Essential Glaucoma Handbook: A guide to assessment and management for eye care professionals*. Malta: Optician

STAMPER, R.L. (1984). The effect of glaucoma on central visual function. *Transactions of the American Ophthalmological Society*, 82, p.792-826.

STARITA, R.J., PILTZ, J.R., LYNN, J.R., FELLMAN, R.L. (1987). Total variance of serial Octopus visual fields in glaucomatous eyes. *Documenta Ophthalmologica*, 49, p.85-90.

STEELE, C. and SPRY, P. (2009). The NICE guideline on diagnosis and management of chronic open-angle glaucoma and ocular hypertension: implications for optometry. *Optometry in Practice*, 10, p.33-50.

STEWART, W.C. and SHIELDS, M.B. (1991). The peripheral visual field in glaucoma: reevaluation in the age of automated perimetry. *Survey of Ophthalmology***,** 36(1), p.59-69.

STEWART, W.C., ROGERS, G.M., CRINKLEY, C.M., CARLSON, A.N. (1995). Effect of cataract extraction on automated fields in chronic open-angle glaucoma. *Archives of Ophthalmology*, 113, p.875-879.

STEWART, W.C., SHIELDS, M.B., OLLIE, A.R. (1988). Peripheral visual field testing by automated kinetic perimetry in glaucoma. *Archives of Ophthalmology*, 106(2), p.202-206.

SULLIVAN-MEE, M., TRAN, M.T.K., PENSYL, D., TSAN, G., KATIYAR, S. (2016). Prevalence, feature, and severity of glaucomatous visual field loss measured with the 10-2 achromatic threshold visual field test. *American Journal of Ophthalmology*, 168, p.40-51.

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

SUSANNA, R., NICOLELA, M.T., SORIANO, D.S., DE CARVALHO, C.A. (1994). Automated perimetry: a study of the Glaucoma Hemifield Test for the detection of early glaucomatous visual field loss. *Journal of Glaucoma*, 3, p.12-16

SUZUMURA, H. (1988). Visual fatigue-like effect in glaucomas with repeated threshold measurement. *Acta Societatis Ophthalmologicae Japonicae***,** 92, p.220-224.

TANNA, A.P., ABRAHAM, C., LAI, J., SHEN, J. (2004). Impact of cataract on the results of frequency-doubling technology perimetry. *Ophthalmology*, 111(8), p.1504-1507.

TERESA, C.M., ANDREA, P., STEFANO, A., VINCENZO, R., MARIA, R.S. (2007). The influence of learning effect on frequency doubling technology perimetry (Matrix). *Journal of Glaucoma*, 16, p.297-301

THAM, Y.C., LI, X., WONG, T.Y., QUIGLEY, H.A., AUNG, T., CHENG, C.Y. (2014). Global prevalence of glaucoma and projections of glaucoma burden through 2040: a systematic review and meta-analysis. *Ophthalmology*, 121(11), p.2081-2090

THE ADVANCED GLAUCOMA INTERVENTION STUDY INVESTIGATORS. (1994). Advanced Glaucoma Intervention Study: 2. Visual field test scoring and reliability. *Ophthalmology,* 101, p.1445-1455.

THE ADVANCED GLAUCOMA INTERVENTION STUDY INVESTIGATORS. (1998). Advanced Glaucoma Intervention Study (AGIS): 4. Comparison of treatment outcomes within race: seven-year results. *Ophthalmology*, 105, p.1146-1164.

THE ADVANCED GLAUCOMA INTERVENTION STUDY INVESTIGATORS. (2000). The Advanced Glaucoma Intervention Study, 6: effect of cataract on visual field and visual acuity. *Archives of Ophthalmology*, 118(12), p.1639-1652.

THE ADVANCED GLAUCOMA INTERVENTION STUDY INVESTIGATORS. (2001). The Advanced Glaucoma Intervention Study: 8. Risk of cataract formation after trabeculectomy. *Archives of Ophthalmology*, 119, p.1771-1779.

THOMPSON, H.S., WALL, M. (2008). Imaging and Perimetry Society (IPS). A History of Perimetry. Available at: **http://webeye.ophth.uiowa.edu/ips/PerimetryHistory**. [Accessed March 27, 2015].

TIELSCH, J.M., SOMMER, A., KATZ, J., ROYALL, R.M., QUIGLEY, H.A., JAVITT, J. (1991). Racial variation in the prevalence of primary open angle glaucoma: the Baltimore Eye Survey. *The Journal of the American Medical Association (JAMA)* 266, p.369-374.

TRAYNIS, I., DE MORAES, C.G., RAZA, A.S., LIEBMANN, J.M., RITCH, R., HOOD, D.C. (2014). The prevalence and nature of glaucomatous defects in the central 10° of the visual field. *JAMA Ophthalmology*, 132(3), p.291-297.

TROPE, G.E. and BRITTON, R. (1987). A comparison of Goldmann and Humphrey automated perimetry in patients with glaucoma. *British Journal of Ophthalmology*, 71, p.489-493.

TSUJI, A., INAZUMI, K., YAMAMOTO, T., KITAZAWA, Y. (1998). Evaluation of the Swedish Interactive Thresholding Algorithm, a new thresholding algorithm, of the Humphrey field analyzer in normal subjevts. *Nihon Ganka Gakkai Zasshi*, 102(6), p.359-364. [Article in Japanese]

TURPIN, A., MCKENDRICK, A.M., JOHNSON, C.A., VINGRYS, A.J. (2002). Performance of efficient test procedures for frequency doubling technology (FDT) perimetry in normal and glaucomatous eyes. *Investigative Ophthalmology and Visual Science*, 43, p.709-715.

TURPIN, A., MCKENDRICK, A.M., JOHNSON, C.A., VINGRYS, A.J. (2003). Properties of perimetric threshold estimates from full threshold, ZEST, and SITA like strategies, as determined by computer simulation. *Investigative Ophthalmology and Visual Science,* 44, p.4787-4795.

VAN DEN BERG, T.J. (1995). Analysis of intraocular straylight, especially in relation to age. *Optometry and Vision Science*, 72, p.115-121.

VAN DEN BERG, T.J.T.P., (RENÉ) VAN RIJN, R., KAPER-BONGERS, L.J., VONHOFF, D.J., VÖLKER-DIEBEN, H.J., GRABNER, G., NISCHLER, C., EMESZ, M., WILHELM, H., GAMER, D., SCHUSTER, A., FRANSSEN, L., DE WIT, G.C., COPPENS, J.E. (2009). Disability glare in the aging eye: assessment and impact on driving. *Journal of Optometry*, 2 (3), p.112-118.

VASHIST, P., TALWAR, B., GOGOI, M., MARAINI, G., CAMPARINI, M., RAVINDRAN, R. D., MURTHY, G.V., FITZPATRICK, K.E., JOHN, N., CHAKRAVARTHY, U., RAVILLA, T.D. , FLETCHER, A. E. (2011). Prevalence of Cataract in an Older Population in India: The India Study of Age-related Eye Disease. *Ophthalmology*, *118*(2-19), 272-278.

VASSILEV, A., IVANOV, I., ZLATKOVA, M.B., ANDERSON, R.S. (2005). Human S-cone vision: relationship between perceptive field and ganglion cell dendritic field. *Journal of Vision*, 5, p.823–833.

VASSILEV, A., MIHAYLOVA, M.S., RACHEVA, K., ZLATKOVA, M., ANDERSON, R.S. (2003). Spatial summation of S-cone ON and OFF signals: effects of retinal eccentricity. *Vision Research*, 43, p.2875–2884.

VESTI, E., JOHNSON, C.A., CHAUHAN, B.C. (2003). Comparison ofdifferent methods for detecting glaucomatous visual fieldprogression. *Investigative Ophthalmology and Visual Science*, 44, p.3873-3879.

VIJAYA, L., ARVIND H., GEORGE, R., BASKARAN, M., RAJU, P., RAMESH, S.V., PAUL, P.G., VIJAYA, L. (2005). Effect of cataract surgery with intraocular lens implant on frequency doubling perimetry. *Current Eye Research*, 30, p.123-128.

VINGRYS, A.J. and PIANTA, M.J. (1999). A new look at threshold estimation algorithms for automated static perimetry. *Optometry and Vision Science*, 76, p.588–595.

VIVELL, P.M., LACHENMAYR, B.J., ZIMMERMAN, P. (1991). Comparative study of various perimetry strategies. *Fortschritte der Ophthalmologie,* 88, p.819–823.

VOLBRECHT, V.J., SHRAGO, E.E., SCHEFRIN, B.E., WERNER, J.S. (2000). Spatial summation in human cone mechanisms from 0 degrees to 20 degrees in the superior retina. *Journal of the Optical Society of America. A, Optics, image science, and vision.* 17, p.641–650.

WABBELS, B.K., REINHARD, O.W., BURK, R.O.W., KOLLING, G. (2001). CLIP: an improved strategy in automated static perimetry. IN: WALL, M. and MILLS, R.P. (eds.) *Perimetry update 2000/2001*. The Hague (The Netherlands): Kugler Publications, p.177-186.

WADOOD, A.C., AZUARA-BLANCO, A., ASPINALL, P., TAGURI, A., KING, A.J. (2002). Sensitivity and specificity of Frequency-doubling Technology, Tendency-oriented Perimetry, and Humphrey Swedish Interactive Threshold Algorithm-fast Perimetry in a glaucoma practice. *American Journal of Ophthalmology*, 133, p.327-332.

WALL, M and JOHNSON, C.A. (2005). Principles and techniques of the examination of the visual sensory system. IN: Miller, N.R., Newman, N.J., Biousse, V., Kerrison, J.B., eds. *Walsh & Hoyt's Clinical Neuro-Ophthalmology*, 6th ed. Vol 1. Philadelphia: Lippincott Willams & Wilkins, p.83-149.

WALL, M., DOYLE, D.K., EDEN, T., ZAMBA, K.D., JOHNSON, C.A. (2013). Size threshold perimetry performs as well as conventional automated perimetry with stimulus size III, V, and VI for glaucomatous loss. *Investigative Ophthalmology and Visual Science*, 54, p.3975-3983.

WALL, M., KUTZKO, K.E., CHAUHAN, B.C. (1997). Variability in patients with glaucomatous visual field damage is reduced using size V stimuli. *Investigative Ophthalmology and Visual Science*, 38, p.426-435.

WALL, M., PUNKE, S.G., STICKNEY, T.L., BRITO, C.F., WITHROW, K.R., KARDON, R.H. (2001). SITA standard in optic neuropathies and hemianopias: a comparison with full threshold testing. *Investigative Ophthalmology and Visual Science*, 42, p.528-537.

WANG, Y and HENSON, DB. (2013). Diagnostic Performance of Visual Field Test Using Subsets of the 24-2 Test Pattern for Early Glaucomatous Field Loss. *Investigative Ophthalmology & Visual Science*, 54, p.756-761.

WEBER, J. (1990). A new strategy for automated perimetry. *Fortschritte der Ophthalmologie ,* 87, p.37–40. [Article in German]

WEBER, J. and KLIMASCHKA, T. (1995). Test time and efficiency of the dynamic strategy in glaucoma perimetry. *German Journal of Ophthalmology*, 4, p. 25-31.

WEIJLAND, A., FANKHAUSER, F., BEBIE, H., FLAMMER, J. (2004). Automated Perimetry. Visual Field Digest. CH-Ko¨ niz: Haag-Streit AG.

WEINREB, R.N. and KHAW, P.T. (2004). Primary open-angle glaucoma. *Lancet*, 363, p.1711-1720.

WERNER, E.B. and BERASKOW, J. (1979). Peripheral nasal field defects in glaucoma. *Ophthalmology*, 86(10), p.1875-1878.

WERNER, E.B., ADELSON, A., KRUPIN, T. (1988). Effect of patient experience on the results of automated perimetry in clinically stable glaucoma patients. *Ophthalmology,* 95, p.764-767.

WERNER, E.B., KRUPIN, T., ADELSON, A., FEITL, M.E. (1990). Effect of patient experience on the results of automated perimetry in glaucoma suspect patients. *Ophthalmology,* 97, p.44-48.

WERNER, E.B., PETRIG, B., KRUPIN, R., BISHOP, K.I. (1989). Variability of automated visual fields in clinically stable glaucoma patients. *Investigative Ophthalmology and Visual Science,* 30, p.1082-1089.

WILD, J.M. (2001). Short wavelength automated perimetry. *Acta Ophthalmologica Scandinavica*. 79, p.546-559.

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

WILD, J.M., KIM, L.S., PACEY, I.E., CUNLIFFE, I.A. (2006). Evidence for a learning effect in short-wavelength automated perimetry. *Ophthalmology*, 113, p.206-215.

WILD, J.M., MOSS, I.D., WHITAKER, D., O'NEILL, E.C. (1995). The statistical interpretation of blue-on-yellow visual field loss. *Investigative Ophthalmology and Visual Science*, 36, p.1398–1410.

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 and Visual Science*, 40(6), p.1152-1161.

WILD, J.M., PACEY, I.E., O'NEILL, E.C., and CUNLIFFE, I.A. (1999a). The SITA perimetric threshold algorithms in glaucoma. *Investigative Ophthalmology and Visual Science*, 40 (9), p.1998-2009.

WILD, J.M., SEARLE, A.E., DENGLER-HARLES, M., O'NEILL, E.C. (1991). Longterm follow-up of baseline learning and fatigue effects in automated perimetry of glaucoma and ocular hypertensive patients. *Acta Ophthalmologica (Copenh)*, 69, p.210-216.

WILDBERGER, H. and ROBERT, Y. (1988). Visual fatigue during prolonged visual field testing in optic neuropathies. *Neuro-Ophthalmology,* 8, p.167-174.

WILENSKY, J.T. and JOONDEPH, B.C. (1984). Variation in visual field measurements with an automated perimeter. *American Journal of Ophthalmology*, 97, p.328-331.

WILENSKY, J.T., MERMELSTEIN, J.R., SIEGEL, H.G. (1986). The use of different-sized stimuli in automated PERIMETRY. *American Journal of Ophthalmol*ogy, 101, p.710-713.

WILSON, M.E. (1970). Invariant features of spatial summation with changing locus in the visual field. *The Journal of Physiology*, 207, p.611-622.

WOOD, J.M., SWANN, P.G., STAVROU, E.P. (2000). Visual fields in glaucoma: a clinical overview. *Clinical and Experimental Optometry*, 83(3), p.128-135.

WOOD, J.M., WILD, J.M., HUSSEY, M.K., CREWS, S.J. (1987). Serial examination of the normal visual field using Octopus automated projection perimetry. Evidence for a learning effect. *Acta Ophthalmologica,* 65, p.326-333.

WOOD, J.M., WILD, J.M., SMERDON, D.L., CREWS, S.J. (1989). Alterations in the shape of the automated perimetry profile arising from cataract. *Graefe's Archive for Clinical and Experimental Ophthalmology*, 227(2), p.157-161.

YAMAGISHI, M., YAMABA, K., KUBO, C., NOKURA, K., NAGATA, M. (2008). Effects of LED lighting characteristics on visual performance of elderly people. *Gerontechnology,* 7, p.243.

YAO, K and FLAMMER, J. (1993). Relationship cataract density and visual field damage. *European Journal of Ophthalmology*, 3(1), p.1-5.

YAQUB, M. (2012). Visual fields interpretation in glaucoma: a focus on static automated perimetry. *Community Eye Health Journal*, 25 (79-80), p.1-8.

YENICE, O and TEMEL, A. (2005). Evaluation of two Humphrey perimetry programs: full threshold and SITA standard testing strategy for learning effect. *European Journal of Ophthalmology*, 15 (2), p.209-212.

YOUNG, W.O., STEWART, W.C., HUNT, H., CROSSWELL, H. (1990). Static threshold variability in the peripheral visual field in normal subjects. *Graefe's Archive for Clinical and Experimental Ophthalmology*, 228, p.454-457.

ZEPPIERI, M., BRUSINI, P., PARISI, L., JOHNSON, C.A., SAMPAOLESI, R., SALVETAT, M.L. (2010). Pulsar perimetry in the diagnosis of early glaucoma. *American Journal of Ophthalmology,* 149(1), p.102-112.

ZHONG, Y., CHEN, L., CHENG, Y., HUANG, P. (2008). Influence of learning effect on blue-on-yellow perimetry. *European Journal of Ophthalmology*, 18, p.392-399.

ZUCKERMAN, J.L., MILLER, D., DYES, W., KELLER, M. (1973). Degradation of vision through a simulated cataract. *Investigative Ophthalmology*, 12, p.213-224.

ZULAUF, M., LEBLANC, R.P., FLAMMER, J. (1994). Normal visual fields measured with Octopus program G1. II. Global visual field indices. *Graefe's Archive for Clinical and Experimental Ophthalmology*, 232(9), p.516-522.

APPENDICES

A3.1: Shapiro-Wilk test for the mean differences of MS between visits according to location using SPARK Precision (see Chapter 3, pg 104)

Location	Statistic	df	p
central	.956	16	0.592
mid peripheral	.909	18	0.082
peripheral	.946	32	0.111

A4.1: Shapiro-Wilk test for VF data of normal subjects (see Chapter 4, pg 120)

Variables	Statistic	df	p
Age	0.890	83	<0.001*
Spherical equivalent	0.939	83	0.001*
MS of SITA Standard	0.990	83	0.775
MS of SPARK Precision	0.957	83	0.008*
Test duration of SITA Standard	0.923	83	<0.001*
Test duration of SPARK Precision	0.981	83	0.276

A5.1: Shapiro-Wilk test for age and spherical equivalent of glaucoma and age-matched control group (see Chapter 5 pg 142)

		Shapiro-Wilk		
Group		Statistic	df	Sig.
Age	Glaucoma	0.965	39	0.264
	Normal	0.963	45	0.162
SE	Glaucoma	0.932	39	0.021*
	Normal	0.918	45	0.004*

***** Data is not normally distributed

A5.2: Shapiro-Wilk test for VF data of glaucoma and age-matched control group (see Chapter 5 pg 142)

			Shapiro-Wilk		
			Statistic	df	Sig.
Mean sensitivity (MS)	SS	Glaucoma	0.828	39	<0.001*
		Normal	0.981	45	0.652
	SP	Glaucoma	0.860	39	<0.001*
		Normal	0.983	45	0.723
Mean deviation (MD)	SS	Glaucoma	0.814	39	<0.001*
		Normal	0.960	45	0.126
	SP	Glaucoma	0.889	39	0.001*
		Normal	0.841	45	<0.001*
Pattern Standard Deviation (PSD)	SS	Glaucoma	0.850	39	<0.001*
		Normal	0.960	45	0.125
	SP	Glaucoma	0.737	39	<0.001*
		Normal	0.976	45	0.469
Number of abnormal pattern deviation points (NAPDP)	SS	Glaucoma	0.896	39	0.002*
		Normal	0.887	45	<0.001*
	SP	Glaucoma	0.799	39	<0.001*
		Normal	0.702	45	<0.001*
Test duration (s)	SS	Glaucoma	0.883	39	0.001*
		Normal	0.934	45	0.014*
	SP	Glaucoma	0.956	39	0.135
		Normal	0.958	45	0.102

* Data is not normally distributed

A 5.3: Shapiro-Wilk test for AGIS score of glaucoma group (see Chapter 5 pg 155)

	Shapiro-Wilk test		
	Statistic	df	p
SITA Standard (SS)	0.768	39	< 0.001
SPARK Precision (SP)	0.651	39	< 0.001

A 5.4: Shapiro-Wilk test for size and depth of glaucomatous field defect

(see Chapter 5 pg 157)

		Shapiro-Wilk test		
		Statistic	df	p
Size	SS	0.903	39	0.003
	SP	0.590	39	<0.001
Depth	SS	0.795	39	<0.001
	SP	0.611	39	<0.001

A 6.1: Shapiro-Wilk test for age and spherical equivalent of cataract patients

(see Chapter 6 pg 179)

		Shapiro-Wilk		
Group		Statistic	df	Sig.
Age	Cataract	0.920	31	0.024*
	Normal	0.937	31	0.066
SE	Cataract	0.964	31	0.368
	Normal	0.858	31	0.001*

***** Data is not normally distributed

A 6.2: Shapiro-Wilk test for VF data of cataract and age-matched control group

(see Chapter 6 pg. 179)

			Shapiro-Wilk		
			Statistic	df	Sig.
Mean sensitivity (MS)	SS	Cataract	0.804	31	<0.001*
		Normal	0.952	31	0.176
	SP	Cataract	0.852	31	0.001*
		Normal	0.968	31	0.454
Mean deviation (MD)	SS	Cataract	0.871	31	0.001*
		Normal	0.894	31	0.005*
	SP	Cataract	0.823	31	<0.001*
		Normal	0.774	31	<0.001*
Pattern Standard Deviation (PSD)	SS	Cataract	0.706	31	<0.001*
		Normal	0.942	31	0.092
	SP	Cataract	0.836	31	<0.001*
		Normal	0.955	31	0.220
Test duration (s)	SS	Cataract	0.941	31	0.091
		Normal	0.960	31	0.299
	SP	Cataract	0.973	31	0.603
		Normal	0.943	31	0.102

* Data is not normally distributed