Survey of Ophthalmology xxx (xxxx) xxx

Contents lists available at ScienceDirect

Survey of Ophthalmology

journal homepage: www.elsevier.com/locate/survophthal

Review article

Visual field testing in glaucoma using the Swedish Interactive Thresholding Algorithm (SITA)

Jeremy C.K. Tan [a,b,c,*], Jithin Yohannan [d], Pradeep Y. Ramulu [d], Michael Kalloniatis [e,f,g], David P. Crabb [b], Jonathan Crowston [c,h], Jack Phu [e,i,j,**]

[a] Faculty of Medicine and Health, University of New South Wales, Kensington, NSW, Australia
[b] Optometry and Visual Sciences, City, University of London, Northampton Square, London, United Kingdom
[c] Save Sight Institute, University of Sydney,Camperdown, NSW, Australia
[d] Wilmer Eye Institute, Johns Hopkins University School of Medicine, Baltimore, MD, USA
[e] School of Optometry and Vision Science, University of New South Wales, Kensington, NSW, Australia
[f] School of Medicine (Optometry), Deakin University, Waurn Ponds, Victoria, Australia
[g] University of Houston College of Optometry, Houston, TX, USA
[h] Singapore Eye Research Institute, DUKE-NUS Singapore, Singapore
[i] Faculty of Medicine and Health, University of Sydney, NSW, Australia
[j] Centre for Eye Health, UNSW, Sydney, NSW, Australia

ARTICLE INFO

Keywords:
visual fields
static automated perimetry
Humphrey Field Analyzer
glaucoma
Swedish interactive thresholding algorithm (SITA)

ABSTRACT

The Swedish Interactive Thresholding Algorithm (SITA) is the main measurement acquisition algorithm used on the Humphrey Field Analyser, the most commonly used instrument for visual field (VF) assessment worldwide. We compare the sensitivity outputs and reliability parameters of the three currently available SITA algorithms—SITA Standard (SS), Fast (SF), and Faster (SFR), with a focus on the newly released SFR and the 24–2C test grid. SFR displays similar sensitivity outputs to SS and SF, but may not be interchangeable with SS in eyes with more severe VF loss. The reliability metric with the greatest impact on VF reliability is the level of false positives, although the recommended 15 % false positive cut off may be inappropriate as a threshold for judging whether a test is reliable and should be included for use in SFR. Finally, the 24–2C grid may be useful in flagging the presence of a clustered central VF defect, while the 10–2 grid can be used to more comprehensively characterize central field defects. We also discuss strategies to improve testing frequency in clinical practice.

1. Introduction

The assessment of visual field (VF) function using standard automated perimetry is an essential component in the care of patients with suspected or manifest glaucoma. VF testing is focused on identifying patients who demonstrate disease progression so that appropriate management strategies can be implemented. The Humphrey Field Analyzer (HFA, Carl Zeiss, Meditec, Dublin, CA) is the most commonly used instruments for VF assessment worldwide.[1] In clinical practice, the Swedish Interactive Thresholding Algorithm (SITA), available in several strategies and across various test grids, is the most commonly deployed algorithm on the HFA.

In recent years, 2 new developments to the HFA have been introduced—the SITA-Faster (SFR) strategy and the 24–2C grid.[2,3] SITA-Faster reduces testing duration substantially compared to older test strategies and produces similar results, and therefore offers potential advantages in clinical workflow. Several studies have, however, demonstrated a difference in sensitivity and reliability outputs that need to be considered when transitioning patients from older SITA strategies (Standard and Fast) to SFR.[4,5] As such, it is important for clinicians to appraise critically the advantages and limitations of SFR when applied to real world practice.

Abbreviations: CIGTS, Collaborative Initial Glaucoma Treatment Study; FL, Fixation losses; FP, False positive; FN, False negative; FT, Full threshold; GT, Gaze tracker; HFA, Humphrey Field Analyser; MD, Mean deviation; PSD, Pattern standard deviation; SF, Sita Fast; SFR, Sita Faster; SITA, Swedish interactive thresholding algorithm; SPE, Seeding point error; SS, Sita Standard; VF, Visual field; VFI, Visual field index.

* Correspondence to: Prince of Wales hospital eye clinic, Level 4, The Prince of Wales Hospital, High Street Building, Randwick, NSW 2031, Australia.
** Corresponding author at: School of Optometry and Vision Science, University of New South Wales, Kensington, NSW, Australia.
E-mail addresses: jeremy.c.tan@unsw.edu.au (J.C.K. Tan), jack.phu@unsw.edu.au (J. Phu).

https://doi.org/10.1016/j.survophthal.2024.09.005
Received 2 July 2024; Received in revised form 28 August 2024; Accepted 12 September 2024
Available online 29 September 2024

We compare the outputs and reliability parameters of the SITA algorithms, with a focus on SFR and the 24–2C test grid. We discuss practical aspects of testing, such as the interpretation of standard reliability metrics, the utility of the 24–2C test grid in relation to 24–2 and 10–2 for central VF testing, and strategies to improve testing frequency in clinical practice.

2. The Swedish Interactive Thresholding Algorithm

The SITA-Standard (SS) strategy was designed to integrate real time estimates of threshold values and threshold error estimates to reduce test times significantly. This returned sensitivity outputs comparable to older algorithms such as full threshold (FT) which used a more traditional staircase reversal thresholding approach. SITA algorithms replaced the latter with a more efficient psychophysical procedure known as a maximum likelihood algorithm, which allowed thresholds to be estimated with fewer stimulus presentations Other innovations included adapting the interstimulus interval to the patient's response speed and novel methods to estimate test reliability without using explicit catch trials.[6] Improvements on the HFA II with faster stepper motors also enabled stimuli to be presented as quickly as patient reaction times would allow. These modifications allowed testing times to be decreased by approximately 50 % compared with FT,[7–14]. The FT approach with the associated long duration would be impractical in modern clinical practice.

Several studies have shown that the sensitivity and specificity in detecting glaucomatous VF defects using SS were excellent (between 96 % and 100 %) when using FT as a reference standard.[10,12,15] Both strategies were, however, not directly interchangeable in longitudinal monitoring. VF defects were found to be wider and shallower, and global sensitivity indices higher in SS compared to FT.[7,9,13,14,16–18] Certain VF endpoint criteria were also found to differ when switching from FT to SS.[19] For instance, while switching from FT to SS has been shown to produce similar MD values in the Collaborative Initial Glaucoma Treatment Study (CIGTS), the mean CIGTS VF score was significantly lower and produced a less abnormal Glaucoma Hemifield Test (GHT) results than FT.[20] In perimetrically naïve individuals, using SS was found to result in more depressed points on the pattern deviation (PD) probability map than FT.[21] The test-retest variability profile of SS was, however, found to be similar,[7] or lower compared with, FT.[13,16,21,22] FT has also been found to yield approximately twice as many false negatives as SS.[23] The reasons for this is unclear, but the shorter duration of SS compared to FT could result in less fatigue and an improved test performance in the former and consequently more consistent outputs.

SITA Fast (SF) was released commercially with SS and was first described in 1998,[8] incorporating further modifications to make testing durations even shorter than SS. In SF, stimulus intensities are altered in 4 dB steps with 1 reversal at all test points except at the 4 primary points;[8] At the 4 primary points (hence forth referred to as seeding points), stimulus sequences are performed until 2 reversals occur.[8] At the remaining test points, stimulus sequences are interrupted at earlier stages than in SS by an increase in the error-related factor cut off value, which decreases the accuracy of test results.[8]

2.1. A new perimetric algorithm on the Humphrey Field Analyzer: SITA-Faster

The 24–2 SITA-Faster paradigm test protocol reduces test duration by approximately 30 % compared to SF and over 50 % compared with SS.[2] Modifications made to SITA-Fast to produce SFR have been comprehensively described by Heijl and colleagues.[2] These include reducing the number of starting stimulus presentations, requiring only 1 staircase test reversal at seeding test points, removal of extra delays in stimulus timing, removal of false negative and blind spot catch trials and removal of second checks at perimetrically blind points (Table 1).[2] Of specific interest, the 25 dB starting stimulus attenuation value for SS and SF (but not SFR) meant that, in cases without significant defects (i.e. healthy individuals or early disease), the test would be starting with a decreased stimulus.

2.2. Sensitivity values using SFR versus SS and SF

SFR has been reported to have similar global sensitivity values to SF, and small sensitivity differences with SS.[2] Heijl and coworkers evaluated 126 patients who underwent SS, SF, and SFR at each of 2 visits and reported mean deviation (MD) to be similar across SS, SF and SFR, while the visual field index (VFI) was significantly lower in SS compared to SF and SFR, although this was clinically negligible.[2] Other studies have also reported similar MD and VFI values between SS and SFR.[25,26] In a study of 74 perimetry-naïve, visually healthy individuals who underwent SS and SFR on the same visit in random order, investigators found no signficant difference in MD, VFI, GHT, foveal threshold and number of depressed points between both strategies.[25] Another study of 49 patients with glaucoma and an average MD of −8.12 dB who underwent SS and SFR on the same visit found excellent correlation for MD and VFI between both strategies (ICC = 98 %).[27]

Subsequent studies, however, demonstrated notable differences in sensitivity indices between SS and SFR in moderate to severe glaucoma.

Table 1
Comparison of test parameters across SITA-Standard, Fast and Faster.[2,6,8,24].

	Full threshold	Standard	Fast	Faster
Starting stimulus intensities at 4 primary (seeding) test points	25 dB at each of 4 seeding test points	25 dB at each of 4 seeding test points	25 dB at each of 4 seeding test points	Age-corrected expected normal threshold level
Number of reversal at primary (seeding) test points	2 reversal sequences	2 reversal sequences	2 reversal sequences	1 reversal sequence
Prior models	Nil; uses 4–2–2 algorithm[6]	Based on distribution of Full threshold normal values	Based on distribution of Full threshold normal values	Based on distribution of SITA-Fast normal values
Retesting at perimetrically "blind" locations (<0 dB)	Present	Two checks	Two checks	One check
Timing windows	Variable time intervals between stimuli, but less compared to SITA	Variable time intervals between stimuli. Extra 300-ms delay after non-seen stimuli	Variable time intervals between stimuli. Extra 300-ms delay after non-seen stimuli	Variable time intervals between stimuli. No delay after non-seen stimuli
Use of false positive catch trials	Present	Nil; estimated based on patient response time measurements	Nil; estimated based on patient response time measurements	Nil; estimated based on patient response time measurements
Use of false negative catch trials	Present	Present	Present	Off by default; present by user selection
Use of fixation loss catch trials	Present	Present	Present	Off by default; replaced by gaze tracker
Test duration in seconds (SD)	870[a]	369.5 (64.5)*	247 (56.7)*	171.9 (45.3)*

[a] Representative mean test duration as reported by Bengtsson and Heijl[7] *Representative mean and SD test durations as reported by Heijl et.al.[2]

These differences are important when transitioning patients to the new SFR strategy in the real world. In a cross-sectional study of 70 eyes of 70 patients with manifest and suspect glaucoma who underwent SS, SF, and SFR in random order at the same visit, investigators reported a mean difference of 1.5 dB between SFR and SF/SS.[28] There was no statistically significant difference in pattern standard deviation (PSD) and VFI across the 3 algorithms.[28] A study of 364 healthy, suspect and glaucomatous eyes that underwent SS and SFR on the same visit found that SFR generally resulted in higher sensitivity values than SS, and SFR sensitivity values tended to be higher, as SS sensitivity values were lower.[5] The differences were more pronounced in patients with glaucoma and in regions with greater visual field loss.[5] In another study of 766 eyes of 421 patients who were transitioned from SS to SFR, Pham and coworkers examined the difference in MD between the SS-SS and SS-SFR sequences.[4] They found no significant difference in MD between SS to SS and SS to SFR sequences in patients with mild or suspect glaucoma, but a 0.87 dB improvement when transitioning from SS to SFR in moderate glaucoma, and 1.49 dB improvement in advanced glaucoma.[4] Using SFR in patients previously tested with SS may therefore conceal disease progression in patients with moderate-severe glaucoma.[4] This was supported by another study of 59 subjects who underwent SS and SFR on the same visit that demonstrated a worse MD in glaucoma patients when using SS (-3.17 versus −2.81 dB).[29] The agreement in sensitivity values between SS and SFR was also worse in severe compared to mild and moderate glaucoma.[29]

An important reason for observed differences in sensitivity outputs between SFR, SF and SS is fatigue as test duration increases, which has been shown to modify contrast sensitivity[30,31] (described later in Section 4.6). Specificity for detecting field defects may therefore decrease with longer test durations. Furthermore, shorter tests may have narrowed statistical limits for normality and consequently similar or better diagnostic sensitivity as a more time-consuming test. The normal values are also different among strategies, although deviations of sensitivity measures from normal age-corrected threshold values in each strategy are similar. While outputs of SF and SFR are therefore often compared to SS, the latter need not necessarily be deemed the gold standard for SAP sensitivity outputs.

Nevertheless, switching between SS and SFR in eyes with greater VF loss should be approached judiciously because of potentially larger differences in output sensitivity or probability score results.[5] In such cases of severe visual loss, the algorithms are not interchangeable.[5] Switching strategies is facilitated by the fact HFA progression event analysis uses empirically-determined cross-strategy significance limits when looking for change in patients whose follow-up tests have been performed using a different strategy than was used in the baseline tests. Switching is also facilitated when using progression rates that are based on VFI values instead of MD values, as cross-strategy differences in VFI are much smaller than in MD.

In an archetypal analysis of VF patterns in eyes that transitioned from SS to SFR, investigators reported an increased likelihood of SFR preserving patterns reflecting a normal VF and lower tendency to preserve archetypal patterns of VF abnormalities compared to consecutive SS exams in the same eye.[32] In other words, normal VF patterns tended to be repeatable, while there were differences in eyes with VF loss. The difference in sensitivity between SF and SFR may be less pronounced; a study of 93 eyes of healthy and glaucoma patients who underwent 24–2 SF and SFR on the same visit found no statistically significant difference in MD, VFI, and number of depressed points in probability plots between both strategies.[33]

In clinical practice, SFR may, however, be preferred by patients with advanced field loss who have little remaining field and may perceive only a few stimuli during the test. These patients are also often older and more frail. In SFR retesting at perimetrically blind locations (locations where the subject has not responded to a stimulus of maximum intensity) is not performed, unlike in older strategies where a second maximum intensity stimulus is presented (Table 1).[2]

2.3. Reliability metrics using SFR versus SS and SF

A practical concern with SFR is the associated increase in rates of unreliable results as described using conventional manufacturer-recommended guidelines such as >15 % false positive rate.[4,5,34] Phu and Kalloniatis previously found the rate of unreliable results to be significantly higher in SFR compared to SS (29.3 % versus 7.7 %) in a cohort of 364 eyes that underwent both strategies on the same day.[5] The biggest contributor to an unreliable result in SFR was false positive rate > 15 %, followed by seeding point errors and > 6 degrees of eye movement > 20 % of the time. Other studies have found a smaller increase in false positive errors with SFR compared to SS. Pham and coworkers[17] found the level of FP errors to be 2 % higher in SFR compared to SS, similar to results published by Heijl and coworkers.[9] The difference between both strategies was, however, noted to diminish with more severe glaucoma, possibly due to the extent and depth field loss offsetting the more difficult test conditions of SFR.[17] This could also explain the higher incidence of results flagged as unreliable in SFR in the Phu and coworkers' study, which contained a higher proportion of subjects with no or early VF loss.[5] The effect of disease stage on VF reliability interpretation remains an important clinical consideration.

Although the utility of using false positive rate as a representation of test reliability has consequently been questioned,[35] recent papers have provided further insight into the subject, given the relatively higher rates found using SFR.[24,36–38] The elimination of a delay after non-seen stimuli in SFR may lead to more true responses recorded as false positives if the patient is slow to respond, which could explain the elevated false positive rates. The impact of FP rates on sensitivity results in the SITA strategies is discussed further below (Section 3: practical aspects of testing). Note that since the Heijl-Krakau blind spot monitor and false negative rates are not reported in SFR, we do not discuss these metrics of reliability.

2.4. Variability of sensitivity outputs of the SITA strategies

Studies that have quantified VF variability have used test-retest methods or regression residuals as a surrogate measure in larger datasets.[39,40] The variability of SITA has been shown to worsen/increase with increasing defect depth, i.e. at lower sensitivity levels, which is a characteristic of the visual system and not from the testing strategy.[2,41,42] Rabiolo and coworkers examined 4747 eyes of 3095 patients from a retrospective cohort tested with SS and found that variability as measured by standard deviation of residuals from pointwise exponential regression varied with threshold sensitivity and eccentricity.[39] Variability ranged from 2.0 dB at a sensitivity of 33 dB to a peak of 5.5 dB at a sensitivity of 11 dB.[39] The test-retest variability of SFR has been evaluated in a large cohort of healthy, glaucoma suspect and glaucomatous eyes. The mean global and pointwise variability was 2.17 ± 1.2 dB and 2.17 ± 2.9 dB, respectively. Pointwise and global variability was observed to increase with worsening threshold sensitivity and was greater for peripheral compared with central test locations.[43] The peak in variability around 10 dB is due to the threshold value approaching the HFA's maximum stimulus intensity at 0 dB. Calculations that correct for data censoring suggest that variability does not peak, but continues to increase rather quickly as visual sensitivity decreases.[44]

Studies have also compared the variability profile of the SITA strategies. A study of 30 patients examined with FT, Fastpac, and SF on each of 2 visits showed lower test-retest variability in SF compared to FT.[8] Other studies have shown that SF displays greater variability than SS and FT.[11,45] In a study of 49 patients who underwent FT, SS and SF at each of 4 visits, SS displayed lower variability than FT, while in areas of low sensitivity (below 25 dB), test-retest variability was greater with SF compared to FT and SS.[45]

Heijl and coworkers performed SFR, SF and SS on 1 eye of 126 glaucoma and glaucoma suspects patients on each of 2 separate clinic visits 1 day to 2 weeks apart. They reported test-retest variability

standard deviation to be below 2 dB at a threshold sensitivity of 30 dB using SFR, which increased to a peak of over 7 dB at a threshold sensitivity of 7 dB before decreasing as threshold sensitivity approached zero. Test-retest variability was found to be small and statistically nonsignificant between the 3 test algorithms.[2]

The normal values are also different among strategies. These test strategies may not, however, have a clinically significant effect on test results. In a large study of 473,252 SS and SF VF tests in over 88,000 patients, although SF was found to be more variable than SS, differences in measurement precision did not result in a meaningful difference in improving time to VF progression.[46] Giammaria and coworkers developed nomograms for the conversion of FT and SF tests to SS, and found that accuracies were negligibly different from test-retest differences with SS.[47]

2.5. Predictors of variability

Age, visual acuity (VA), mean total deviation, and false negatives have been found to be associated with test-retest variability in SS and SF with VA worse than 6/18 (20/60) Snellen equivalent associated with increased variability.[48] Rabiolo and coworkers found Asian descent, higher false negative and positive rates, worse baseline MD, longer follow-up to be predictors of global visual field fluctuation.[49] They also reported that predictors of greater pointwise variability include worse baseline MD, greater false negative and positive rates, faster VF decay, longer follow up and higher VF frequency.[49] In a retrospective cohort study of 1531 eyes from the Duke Glaucoma Registry, the authors found VF variability as measured by standard deviation from linear regression trend lines to be 1.26 dB for White and 1.53 dB for Black patients.[40]

Other factors that affect variability include stimulus size,[50,51] technician experience, time of day, season, and the percentage of false-positive answers.[52] In our study of test-retest variability of SFR, we found that worse baseline MD and abnormally high sensitivity on GHT were significantly associated with increased variability.[43]

Overall, the limited evidence to date demonstrates that test-retest variability across the 3 SITA algorithms is comparable. In the presence of factors that increase variability, repeat/increased testing may be beneficial in improving the estimation of true sensitivity change.

In summary, SFR displays similar sensitivity outputs to SS and SF in healthy eyes and early glaucoma, it may not be interchangeable with SS in eyes with more severe field loss because of potentially larger differences in output sensitivity or probability score results. Further, limited studies to date show that SFR displays similar test-retest variability profile to SS and SF. Finally, increased age, Asian descent, worse visual acuity, worse baseline MD, and abnormal reliability indices (false positives, false negatives and glaucoma hemifield test) may be associated with increased variability.

3. Visual field test grids on HFA using SITA

The 24–2 grid on the HFA tests 54 evenly-spaced locations 6 degrees apart, comprising 52 within the central 24 degrees from fixation and 2 nasal locations near 30 degrees.[53] Within the central 10 degrees from fixation are 4 inner central points and 8 paracentral points. The 10–2 grid tests 68 evenly spaced locations 2 degrees apart within the central 10 degrees from fixation. The SFR strategy is only available on the 24–2 grid, while the SS and SF strategies are available on both. Importantly, early studies have suggested that the 24–2 grid may miss early central defects.[54] Therefore, the recommendation from those studies was that some patients may benefit from early central visual field testing using the 10–2. This notion has been challenged by later studies.[55] Optical coherence tomography facilitates high resolution structural evaluation of the optic nerve and macular ganglion cell layer, identifying potentially preperimetric cases of glaucoma. Nonetheless, measurement of central visual field defects remains important for understanding functional impact and its progression.[53] In glaucomatous eyes with parafoveal scotomata, the 10–2 grid also detects more progressing eyes than the 24–2 over time, warranting closer surveillance of the central VF using the 10–2 in patients with this specific VF defect.[56] The studies above reinforce the importance of characterising the pattern of VF defect, as the relative utility of 10–2 and 24–2 test grids for the individual patient can then be determined. This concept relates to both test location selection as well as dynamic range. Progression of primarily peripherally-located VF defects, such as nasal steps or high arcuate defects, would be best sampled using a 24–2 test grid. Conversely, primarily centrally-located defects could be better assessed using a 10–2 test grid.

Notably, recent work by Phu and coworkers has contributed to the understanding of when to deploy a 10–2 test grid for central visual field testing by considering the dynamic range of the test.[57] This phenomenon was termed the "functional vulnerability zone". In brief, the 10–2 test grid contains test locations that occupy spaces between 24–2 test locations within the central visual field. Accordingly, these additional test locations present an additional "dynamic range" for testing, as they potentially afford more opportunities to detect progressive defects at the border of scotomata; however, the benefits of additional test locations are not always present, as 10–2 test locations that exhibit deep defects do not meaningfully contribute to progression measurement. Thus, they proposed parameters on the 24–2 that can guide clinicians on the utility of performing a 10–2 test using the functional vulnerability framework.[57] Testing using both 10–2 and 24–2 grids in a single clinic visit is, however, often impractical because of time and other resource constraints.[58] Modifications to existing test grids have been suggested to address this, such as adding test points to the 24–2.[59,60] The proposals by Ehrlich and coworkers and Chen and coworkers are, however, not available in clinical practice and not commonly used. In a modelling study, Rafla and coworkers have recently proposed a critical number of test locations within the central VF test grid that can optimize structure-function concordance (8–14 test locations), which are not necessarily location-specific.[61] The 24–2C is a recent test grid on the HFA that incorporates 10 additional test points within the central 10 degrees from fixation—5 each in the superior and inferior hemifields (within the "critical number" determined from our modelling study). The points are asymmetrically distributed and derived from the 10–2,

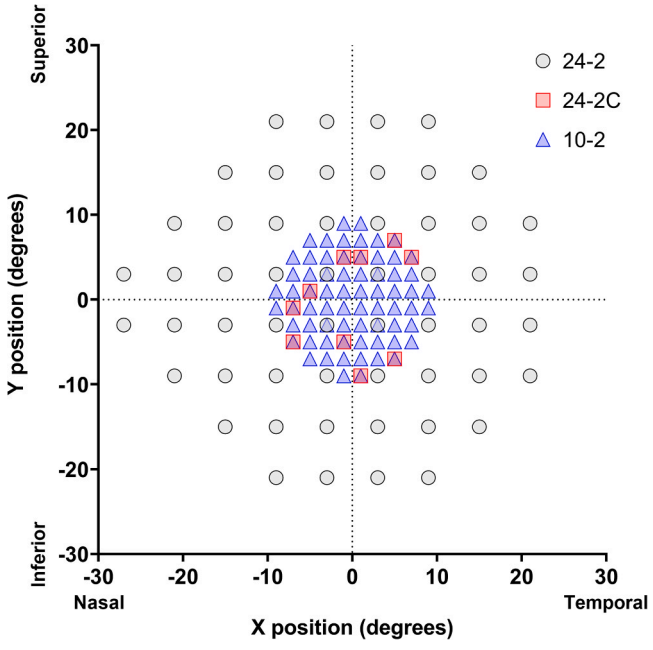

Fig. 1. Test point locations of the 24–2, 24–2C and 10–2 grids superimposed on the same map. The 24–2C comprises the central red points and the 24–2 grid points. The visual field is in a right eye format.

and were purportedly derived from test locations commonly affected by glaucoma (Fig. 1).[3] Sensitivity measurements on the 24–2C are obtained via the SFR strategy. Given that a hybridized grid such as the 24–2C could potentially offer advantages for detecting defects in both peripheral and central visual field testing, several studies have compared the perimetric outputs between the 24–2C and the clinical standard 24–2 and 10–2 grids.

3.1. Comparison of 24-2C and 24-2 test grids

In the first independent study on the 24–2C, Phu and Kalloniatis compared global sensitivity indices and structure-function concordance in 100 eyes of 100 patients with suspected and manifest glaucoma who underwent 24–2 (SS and SFR) and 24–2C SFR on the same visit.[3] The average MD and central mean sensitivity was found to be worse for the 24–2C SFR compared to 24–2 SS, but no different between 24–2C SFR and 24–2 SFR.[3] The former difference was, however, not likely to be clinically significant.[3] For the majority of patients, the additional test points in the 24–2C did not add further statistically significant clusters (3 or more contiguous points) of sensitivity reductions compared to the 24–2 across the entirety of the 24–2 grid. Within the central 10 degrees, however, there was a tendency for the 24–2C to identify more clusters of sensitivity reduction compared to the 24–2 test grid alone.[3] The 24–2C therefore offers the potential advantage of alerting the clinician to the presence of central visual field loss simply by virtue of having a greater number of test locations.[3] Behera and coworkers examined 60 eyes of glaucoma patients who had a confirmed glaucomatous field defect on 24–2 SS on a previous visit.[62] Each eye then underwent a 24–2C and 10–2 test on the same visit. They found that the 24–2C detected 5.5 and 2 more defective points than the 24–2 on the central 10 degrees of the total deviation and pattern deviation plot, respectively. Excellent correlation in MD and PSD was found between 24–2C and 24–2.[62] Test duration for the 24–2C was found to be a median 26 seconds longer in the 24–2C compared to 24–2 in this cohort of healthy, suspect, and early glaucoma eyes.[3]

The greater number of points in the central 10 degrees of the 24–2C grid has enabled greater macular structure-function associations compared to the 24–2 grid.[3,63] In 150 eyes of healthy, preperimetric and perimetric glaucoma subjects, there were significantly stronger structure-function correlations calculated by weighted correlation coefficients between the average and sectoral macular ganglion cell-inner plexiform layer thickness and central VF mean sensitivity using the 24–2C grid compared to the 24–2 grid. Phu and Kalloniatis also found a greater number of test locations exhibiting structure-function concordance in the central visual field when using the 24–2C test grid compared to the 24–2 test grid.[3] Overall, these studies suggest that the 24–2C identifies more central VF defects and instances of structure-function concordance compared to the 24–2. Strong correlation in global indices however suggest similar depth of defect information obtained using both test grids.

3.2. Comparison of 24–2C and 10–2 test grids

Phu and Kalloniatis compared global indices, VF defects and structure-function correlations with ganglion cell layer thickness in 131 glaucoma and 57 glaucoma suspect patients who underwent 24–2C SFR and 10–2 SF in random order.[64] The 24–2C and 10–2 test grids returned similar MD, PSD and central mean sensitivity and proportionally similar amounts of central visual field loss.[64] The additional points in the 10–2 grid returned more "clusters" of defects defined as a pair of contiguous points both at $P < 0.01$ level, or at least 3 points at $P < 0.05$ level with at least 1 point at $P < 0.01$ level.[64] Importantly, it was noted that typical "contiguity" criteria for clusters of visual field defects could not be applied to the 24–2C due to the asymmetric distribution of test locations. There was also a greater rate of structure-function concordance with ganglion cell layer thickness compared with the 24–2C test grid. Finally, test duration for the 10–2 SF was a median 47 s longer than the 24–2C.[64]

As noted above, Behera and coworkers's study found 2.5–3 times the number of defective points on the total deviation and pattern deviation plots respectively when using the 10–2 compared to the 24–2C There was a high intraclass correlation of 0.80 for MD and PSD between the 24–2C and 10–2.[62] In combination, the studies described above suggest a dose-dependent effect of the number of test locations on parameters important for VF interpretation, as discussed by Rafla and coworkers[61] Therefore, while the 24–2C may therefore be useful in flagging the presence of a clustered central visual field defect, the 10–2 can be used to more comprehensively characterise a central field defect.[64]

In summary, the 24-2C is a recent test grid on the HFA that incorporates 10 additional test points within the central 10 degrees from fixation—5 each in the superior and inferior hemifields. These points are asymmetrically distributed and derived from the 10-2. The 10-2 SF, 24-2C SFR and 24-2 SFR have been shown to produce similar global sensitivity and central mean sensitivity outputs. While the 24-2C may be useful in flagging the presence of a clustered central visual field defect, the 10-2 should be used to more comprehensively characterise a central field defect.

4. Practical aspects of visual field testing

4.1. Standard reliability indices—false positive, false negative and fixation losses

Computerised perimetry was initially implemented with 3 parameters to help users evaluate the reliability and usefulness of test results: fixation losses (FL), false negative rate (FN) and false positive rate (FP).[24] FL responses were initially obtained via the presentation of test stimuli at the expected location of the physiologic blind spot (the "Heijl-Krakau" method). This method has several shortcomings though, especially when the blind spot is not situated at the assumed location or if it were improperly mapped out in the initial phases of the test, or if patients move back or tilt their head during testing. It was, however useful at the time of use in the competer perimeter, where the perimetrist could not see the patient's eye. False negative (FN) responses are obtained by presenting suprathreshold stimuli (e.g. 6 dB above a threshold established earlier in the test), and act as a surrogate for inattention. FN responses may, however, be influenced more by the level of visual field damage than on patient vigilance,[65,66] where extensive visual field loss increases the uncertainty of whether a stimulus was actually seen and thus raises the likelihood of spurious FN catch trial activity.[66] A recent study, however, showed that false negative response rate was associated with improved classification of glaucoma, with no association found with false positive rate or fixation losses.[67] Finally, false positive (FP) responses are intended to identify "trigger-happy" behaviour and were originally measured using catch trials in a response window in which no stimulus was presented. Modern SITA strategies, however, estimate FP based on the detection of patient responses during intervals where it is impossible or unlikely for a stimulus to have been seen.[24] The FP rates estimated in the SITA family of algorithms are based on a maximum likelihood method[68] and were based on the distribution of FP results, rather than necessarily representing true reliability or a relationship with sensitivity outputs.[24] The utility of using FP rate as a representation of test reliability has consequently been questioned in recent papers.[24,36–38]

While the importance of FL and FN in assessing test reliability has declined over time, FP responses remain widely used in modern perimetric strategies, including SFR.[24]. Parallel to this, cut-off values for FP rates have also evolved over the past 3 decades, with differences in acceptable limits varying across major clinical trials. For instance, the Ocular Hypertension Treatment Study had an acceptable FP cutoff of 33 %, while subsequent trials like the United Kingdom Glaucoma Treatment Study adopted a lower cutoff of 15 %.[69,70] Currently, the 15 % FP threshold is often used as a criterion for reliability in perimetric

testing. Notably, this is a recommendation by the manufacturer of the HFA.

4.2. Impact of standard reliability indices on mean deviation

FP and FN values are included in the frequency of seeing model for calculating threshold values in SITA tests, and, therefore, higher rates of FP and FN answers influence threshold/sensitivity values less in SITA than in FT and Fastpac tests. Yohannan and coworkers examined the impact of standard reliability indices on SS visual field reliability.[71,72] They calculated predicted MD by multilevel modelling of longitudinal data, and the difference between predicted and observed MD was used as the reliability measure. FP was found to have the greatest impact on VF reliability, and this was most prominent in eyes with moderate or severe disease. For instance, 14 % FP is expected to produce an observed MD 1 dB above predicted MD in moderate and severe glaucoma, compared with 21 % FP required to produce the same change in mild disease.[72] FN had a smaller impact on VF reliability than FP, with a greater percentage of FN generally associated with an observed MD more negative than that predicted by our models.[72] This aligns with previous suggestions of the relationship between visual field defect and FN rate. FL produced almost no impact on MD difference.[72]

Relevant to deployment of SFR, a recent study of 126 patients who underwent SS, SF and SFR at each of two visits examined inter-visit differences in MD and VFI as a function of FP rates.[24] This study found that higher FP rates were associated with greater increases in MD, but the effects were small and dependent on test strategy. For instance, each percentage point increase in FP was associated with an increase in MD of 0.06 dB in eyes with FP $\leq$ 15 %, and 0.04 dB in eyes with FP $>$ 15 %. The positive effect of FP rates on MD was larger with SFR compared to SS and SF, and more pronounced in severe compared to early and moderate glaucoma. The effects of FP on VFI were weaker than that with MD. PD probability maps were also not influenced by changes in FP. The authors concluded that test results should not be discarded solely based on FP response rates.[24] Instead, they proposed a method of evaluating reliability using general height (difference between numerical total deviation values and pattern deviation values), which was found to display stronger relationships with inter visit differences in MD and VFI compared to false positive rates.[24] This metric is, however, currently not yet incorporated in clinical practice.

Tan and coworkers previously calculated the difference in test-retest mean deviation as a function of the false positive rate when using SFR.[43] They observed a gradual increase in MD difference with an increasing FP rate. There was, however, minimal change in MD difference around the recommended 15 % FP cut off, suggesting that this threshold may be inappropriate as a threshold for judging test reliability and data inclusion in SFR. It has also previously been shown that patients with FPs above 15 % returned global indices that were comparable to a result that met conventional reliability criteria.[37] Depending on the acceptable range of sensitivity differences, different inflection points of FP rates could be used with the Tan and coworkers,[43] study showing a delta MD value of 1 dB at around a FP rate of 33 %. Overall, the degree of test reliability captured by standard reliability indices is generally small, with FP responses likely to be the most useful metric.

4.3. Gaze tracker outputs

Gaze tracking (GT) measures degrees of gaze deviation during stimulus presentation and reports GT failures in real time.[36] SFR has been designed to use gaze tracking instead of FL catch trials as a parameter for test reliability. GT outputs are displayed as upward bars that are representative of gaze deviations, with increasing bar sizes indicative of larger gaze deviations, while downward bars indicate failure to capture a signal.[36] Camp and coworkers examined the association between gaze tracker metrics and standard HFA reliability metrics (FL, FP and FN).[36] Gaze tracker metrics were calculated as the percentage of stimuli with gaze deviations between 1 and 2 degrees, 3–5 degrees, and 6 degrees or more, and percentage of stimuli with tracking failure. Although there was a statistically significant association between fixation losses and FNs and GT metrics, the area under the curve calculations and low correlation coefficients indicate that clinical significance was low. They therefore concluded there was no clinically significant correlation between standard reliability metrics and gaze tracker metrics.[36]

Phu and Kalloniatis examined the correlation of GT outputs on the SFR with sensitivity and reliability indices.[38] GT outputs were aggregated into total ticks, sum of amplitudes and average amplitudes. There was a weak correlation between eye movements and MD which was driven by more severe MD values.[38] There was no significant correlation between gaze tracker outputs and false positive rate.[38] Another study found GT parameters to be associated with lower MD values.[73] GT outputs have however been found to improve the detection of progression, with mean progression rates more reliable when stricter gaze tracking reliability criteria is used.[73,74] GT outputs have also been found to be associated with test-retest variability in SS.[75]

Overall, standard reliability and GT metrics tend to be influenced by the severity of field loss,[36] and more importantly, the methodology by which they are determined. Test results should not be discarded solely based on deviation of these parameters outside of recommended reference limits. Visual field tests may also not be truly reliable despite “normal” reliability indices.[71]

4.4. Seeding point errors

Seeding point errors (SPEs) contribute to a large proportion of unreliable visual field results, especially when using SFR[5,76]. They are typically identified at the end of the test by the presence of unusually reduced or elevated sensitivity results at any of the 4 seeding points in the test grid. Phu and Kalloniatis observed a greater occurrence of SPEs in SFR compared with SS, which contributes substantially to the rate of low test reliability especially in SFR.[5] They proposed that the uniqueness of this artefact to SFR (in comparison to SS) was due to the modification to seeding point testing in the SFR strategy.[5] In SFR, the luminance level of the stimuli initially presented at the 4 seeding locations is near the age-expected threshold.[2] This contrasts with SS and SFR which present an initial stimuli at the 25 dB level, which is brighter than threshold in normal eyes. The presentation of stimulus intensities near threshold in SFR may introduce stimulus uncertainty, which consequently alters the response criterion such that a more intense stimulus is required before the subject indicates a response.[76] The second modification with seed point testing in SFR is the use of only 1 reversal for the staircase, unlike the 2 used in SS.[2]

Phu and Kalloniatis proposed a model for detecting SPEs early in the test process to guide to retesting.[76] The model utilises sensitivity results from 9 locations (each of the 4 seeding point locations plus 5 adjacent points) and distinguish SPEs at close to 90 % accuracy. Specifically, if the seeding point sensitivity result is $\leq$26 dB or if the seeding point sensitivity result is different from its neighbouring points by a total of $\geq$ 11 dB, this should alert the perimetrist to a potential need to restart the test.[76] Later, Phu and Kalloniatis examined the impact of SPE on overall visual field indices.[76] They concluded that the presence of SPE (in the context of an otherwise “reliable” result) exerted no substantial effect on the returned global indices. Only the affected seeding points demonstrated spurious defects, thus potentially reducing their effectiveness as a baseline for interpolating adjacent sensitivity results. Thus, this characteristic of SFR remains important to recognise as apparent defects in the seeding point locations could be mistaken for pathologic results.

4.5. Learning effect

Paradigms in perimetry aim to estimate a subject’s contrast sensitivity threshold, i.e. the stimulus contrast at which it will be detected

50 % of the time. One of the largest sources of variability in perimetric testing is instability in the patient's decision rule, i.e. 'how certain they have to be that they saw the visual stimulus before responding?'.[77] The effect of the decision rule is conceptualized by signal detection theory, which describes the subject's internal criterion for response, and can be affected by intrinsic factors such as procedural experience, practice and attention, and extrinsic influences, such as distractors.[78] The learning effect has been frequently examined in the literature as a significant source of test variability.[77] Perimetric learning can be demonstrated in both the short- and long-term and may vary depending on perimetric algorithm.[79–82] Threshold sensitivity has been shown to increase in the first few sessions of VF testing in normal subjects.[81] In 25 glaucomatous subjects who underwent repeated 30–2 FT on 5 sessions at 1 week intervals, mean MD significantly improved between the first and second test session, but not between the second and fifth session.[80] These effects were more pronounced for peripheral than central points and for better than more abnormal points.[80]

In another study of 55 patients who underwent SS twice in a single day, the threshold sensitivity increased on the second examination, with changes more evident in peripheral points.[83] Several early VF tests are required to establish a baseline of visual field function. The learning effect may also be prolonged over a number of years; a study of 160 eyes of 80 patients who underwent standard automated perimetry found that mean sensitivity increased by 0.5 dB over the first year and showed no significant change until after 5 tests despite an expected decline in sensitivity due to aging and disease progression during this time.[77] The intervals between visits is also expected to have an influence on the learning effect, as longer intervals may result in subjects needing to "relearn" how to perform the VF test. Having the opportunity to undergo a trial sequence before the formal test, or simple a second VF test on the same visit, may therefore provide a way of improving perimetric quality (discussed later).

4.6. Test fatigue

Test fatigue may contribute to sensitivity loss as testing progresses. Studies of fatigue in perimetry have shown that contrast sensitivity decreases for most individuals as test time increases where testing was performed in uninterrupted sessions lasting around 30 minutes. The resultant increase in threshold is generally small for normal individuals, but can be marked (exceeding 10 dB) especially in glaucoma.[30,31] Test locations in a relative scotoma or adjacent to it can also display a pronounced deterioration of threshold compared to normal points located away from field defects.[31] Test fatigue may also be influenced by testing order, with the second eye expected to potentially display greater signs of fatigue than the first. In a study of 47 patients with suspected or manifest glaucoma experienced with perimetry, order of eye testing was not found to have any significant effect on MD or reliability indices.[84] Starting with the right eye consistently may nevertheless help keep any perimetric fatigue effects constant over time.[85] In a larger retrospective study of 6901 subjects that underwent at 6 SS or SF tests in routine clinics, investigators assessed fatigue based on test variability by linear regression of MD values. They reported a significant increase in perimetric fatigue effects in the second eye tested, although these effects tended to be small and highly variable among patients.[85] In a large cohort of patients who underwent 2 SFR tests in each eye on the same visit, test duration as a surrogate marker for fatigue was not significantly greater on the second test.[86] This could be due to the shortened duration of SFR, which along with SF may be advantageous in patients where long testing times are prohibitive. A shorter test may not translate to an easier test; for instance, SF presents stimuli closer to thresholds in order to reduce test time, which may make stimuli harder to perceive.[46]

In summary, studies of SS and SF have shown that FP has the greatest impact on VF reliability followed by FN, while FL has minimal impact on reliability. FN and FL are no longer routinely incorporated in the default SFR strategy. In SFR, the recommended 15% false positive cut off may be inappropriate as a threshold for judging test reliability and inclusion/exclusion of an "unreliable test" and requires further evaluation to define a suitable cutoff. SFR can also result in a higher number of tests with poor reliability compared to SS due to seeding point errors – a major artefact observed in this algorithm due to its thresholding properties. Finally, there is minimal association between standard reliability metrics and gaze tracker metrics, which tend to instead be influenced by the severity of field loss.

5. Strategies to increase frequency of testing for earlier detection of progression

Distinguishing VF progression from inherent fluctuations of threshold estimates between tests represents a major clinical challenge in glaucoma.[87] Fluctuations in perimetric performance may mask true progression, especially in individuals prone to higher rates of test-retest variability and also as glaucoma advances.[88] An adequate number of tests is therefore required to achieve sufficient statistical power to confidently document stability or change in visual field function.[89] For instance, it is estimated that at least 6 VF tests need to be performed within the first 2 years to identify progressors at MD rates of −1 dB per year or worse if the subjects exhibit low or moderate levels of perimetric variability.[90] The European Glaucoma Society recommends 3 VF tests per year for the first 2 years following diagnosis to identify rapidly progressing patients.[91] Not only can increased VF testing identify progression sooner, it may also be cost-effective.[92] Many clinicians nevertheless believe these recommendations are difficult to implement with current resources.[92] The frequency of VF examinations is also often insufficient. A multicenter audit of hospital-based glaucoma clinics in England revealed that most patients only receive one VF test per year.[58] Another study of a nationwide claims database on data from 2008 to 2017 in the United States found that more than 75 % of patients with open-angle glaucoma failed to receive at least one test per year.[93] The burden of time and labor to both staff and patients incurred by static automatic perimetry algorithms such as SS are likely contributors to this.[93] Recommendations of the optimal frequency of testing to detect progression have been described elsewhere.[90,94] Stratification of patients into those at low versus high risk of progression based on early visual field, optical coherence tomography, and clinical data may help identify subjects who may benefit from increased testing while balancing limited clinic resources.[95,96]

5.1. Clustered visual field testing to improve detection of change

Clustered measurement strategies in which multiple tests are performed close together have been demonstrated to improve the likelihood of detecting progression compared to evenly spaced testing,[89,97] and improved the efficiency of testing and reduced the number of visits required[98] as modelled by Chauhan and colleagues.[90] The concept of essentially "signal averaging" facilitates a smoothing of variance across results.

A clustered test approach was integrated into the protocol for the United Kingdom Glaucoma Treatment Study, a randomized-controlled trial assessing vision preservation in latanoprost versus placebo.[99] The trial protocol comprised 12 visits (11+ 1 training visit) over 24 months, with 2 SS tests on a training visit 0, visit 1 at month 0, visit 2 at month 2, visit 7 at month 16, visit 8a and 8b at month 18 and visit 11 at month 24.[100] Single VF testing per eye was performed at all other visits spaced evenly around 2–3 months apart. Clustering of tests was intended to increase the precision of estimate of the rate of VF change within the short trial duration of 2 years, and the authors subsequently reported longer VF preservation in the latanoprost versus placebo group.[99]

More recently, Bradley and coworkers demonstrated that a clustered measurement strategy with approximately half the number of total tests performed at either end of a 2-year period improved the accuracy of detecting glaucoma worsening using peripapillary optical coherence

tomography scans[98] and VF testing[97] and reduced the number of tests required compared with evenly spaced measurements. A clustered strategy with a long interval (eg. 2 years) between multiple VF tests at either end may, however, miss intervening rapid progression and is likely inappropriate for deployment in clinical practice. Furthermore, the longer test durations of older algorithms like FT and SS may make intravisit clustering of tests excessively time consuming and likely impractical within most clinical settings other than in clinical trials.

5.2. *Intra-visit clustering of SFR—"frontloading" visual fields*

To address the impracticalities of clustering with older strategies (SS, FT), Phu and Kalloniatis recently proposed a method of intra-visit clustering using 2 (instead of 3) SFR VF tests per eye per visit.[101] This approach, initially termed "frontloading" VFs, leverages the shortened test duration of SFR, which facilitates more frequent testing within the same duration of older strategies. The frontloading protocol took less than 20 minutes to perform (a total of 4 visual field tests) in over 80 % of subjects including rest breaks,[101] and has been found to be well tolerated by patients and technicians.[15] This is in line with a survey of adult patients who regularly undergo VF tests demonstrating patients are prepared to increase the number and frequency of visual tests to obtain more information about their visual status.[90] An example of a simulated patient is shown in Fig. 2, extrapolated from Wang and coworkers.[102] The example of frontloading provided in Fig. 2B has the same number of VF test results as the wait and see approach (Fig. 2A) at 2 years of follow-up (6 tests in total), with a comparable 95 % confidence band and approximation of the ground truth progression rate. Both wait-and-see and frontloading clustering methods are more likely to identify statistically significant progression with narrower confidence bands compared to the current clinical standard.

Frontloaded visual field tests display strong correlation of global sensitivity indices, with a significant improvement of reliability indices from the first to second test.[86,101] Phu and Kalloniatis performed a simulation study of patients with early or suspected glaucoma undergoing different number of visual field tests per visit and the detection of MD change over time.[103] They found that 2 visual field tests per session compared with 1 provided higher case detection rates at 2 years (99 %-99.8 % versus 34.7 %-76.3 %, respectively), a reduced time to detection of progression (3 or 4 visits versus 6–10, respectively), and less

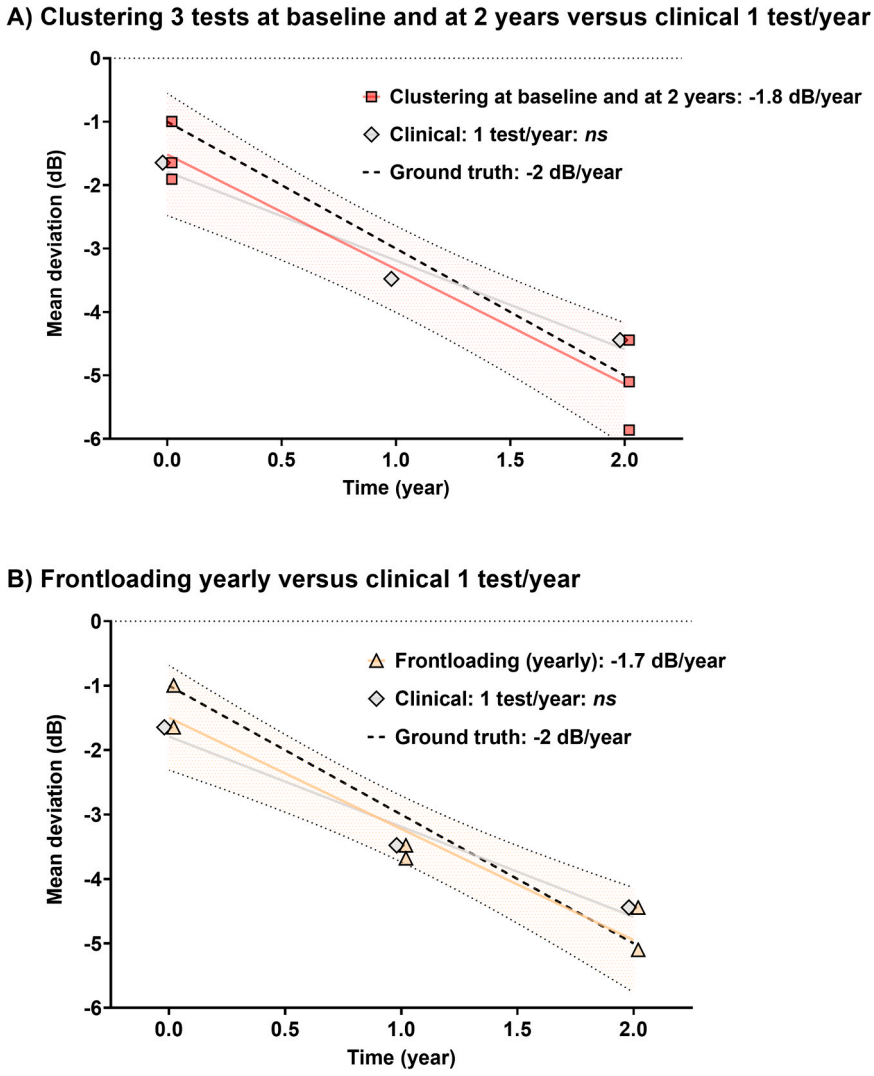

Fig. 2. Examples of visual field (VF) testing approaches in a simulated patient with glaucoma. Symbols represent the MD value of each VF test. The ground truth progression rate is −2 dB/year and shown by the black dashed line. The simulation of one visual field per year does not reach significance (ns) for progression when three visual fields are performed over the two year period (open diamonds). The regression lines and shaded 95 % confidence band of two types of clustered VF test approaches are shown: (A) Three visual field tests at baseline and at 2 years using a clustering approach (red squares) provides a progression rate of −1.8 dB/year, close to the ground truth rate. (B) The frontloading (2 VF per eye per visit, teal triangles) performed yearly to yield a total of six visual fields over the two year period with a predicted progression rate of −1.7 dB/year, also close to the ground truth rate.

negative mean deviation value (-4 dB versus −10 dB, respectively) at the point of mean deviation change identification, especially in the context of unreliable results.[103] These results were also apparent when examining "slow" progressors, which would be the majority of patients expectedly seen in most clinical practices.[104] The above 2 studies examined cross-sections of simulated patients with fixed progression rates. The frontloading protocol was then applied to MD distribution and progression rate data derived from 2 cohorts of patients followed under routine glaucoma clinical care (a Swedish cohort collected by Heijl and coworkers[105] and a Canadian cohort described by Chauhan and coworkers[106]). These contrasting clinical cohorts presented an opportunity to examine the effects of frontloading in more "real world" scenarios with a variety of slow, average, and fast progressors. The inestigators showed that the frontloaded strategy detected more progressors than the non-frontloaded protocol.[102] The time required to detect 50 % of progressors was also 1–1.5 years less using the frontloaded protocol.[102]

Frontloading VFs can provide additional perimetric data to not only compare global sensitivity values, but also to evaluate the consistency of pointwise defects.[107] Tan and coworkers examined pointwise deviation map probability data from 3 sequential tests in a cohort of eyes with glaucoma—each patient had received one SS test on a previous visit, followed by a frontloaded pair of SFR tests at the next visit, to undertake an event-based analysis of pointwise changes seen on the first SFR test. When using only the first SFR test, 73.8 % of the points on SFR 1 appeared to confirm locations of existing defects or non-defects on the pattern deviation grid of the SS test. The remaining points were divided between a possible reversal of a defect or a new defect. The second SFR test allowed the distinction between points on SFR 1 which were repeatable versus nonrepeatable, to provide greater certainty of the consistency of a new defect or reversal of a defect seen on SFR 1.[107]

While test fatigue is a valid concern with intra-visit clustered testing, the authors did not observe a significant decline in performance in the form of worsening false positive rate or an increase in test duration in the second test.[86] While a significant worsening of MD values on the second test in some patients with MD values towards the normal end of the reference range has been observed, this is more likely attributable to a decrease in "trigger-happy behavior" as opposed to fatigue, given that rates of false positive and abnormally high sensitivity responses similarly declined.[86] In a cohort of patients with glaucoma and average MD of −5.83 dB, there was no significant difference in MD between frontloaded tests.[107] A written questionnaire survey of patients who underwent frontloaded testing found that the vast majority of patients reported being comfortable during the test and preferred to complete the tests at a single visit rather than returning for retesting.[108]

Overall, these recent studies demonstrate the potential benefits of a frontloading approach, facilitated by SFR. Moving forward, frontloading as verbiage could be reconsidered to improve real-world uptake, as it implies only performing multiple tests at baseline (Instead, intravisit clustering or multiplex testing have been suggested). Technical challenges of incorporating multiple tests on the same day into the machine's progression algorithms and costs/billing of extra VF tests are, however, important issues to be considered.

5.3. Home monitoring of visual fields using portable devices

Other recent methods to increase perimetric data include home monitoring using portable tablet-based or virtual reality devices have demonstrated good concordance with clinic-based automated perimetry in terms of global sensitivity values,[109–111] with good compliance.[110] In a study of 101 participants with suspected or manifest glaucoma, each participant was given a tablet device for home monitoring of visual fields and advised to perform 6 examinations at weekly intervals;[112] 88 % completed at least 1 home examination and 69 % complete all 6 examinations, indicating overall good short-term compliance.[112] Pointwise sensitivity comparisons between portable devices and clinic-based machines have, however, highlighted notable deviations, with substantial differences based on location. Integrating data from these portable devices with clinic-based devices to detect visual field change also poses another important technical consideration. Discussion of portable perimetry is outside the scope of this review and has been described elsewhere.[113,114]

In summary, the frequency of visual field testing in hospital-based eye services are substantially below recommended minimum testing levels, which limits the detection of visual field progression. Further, shortened test durations of newer strategies such as SFR and SF may facilitate more frequent testing to meet recommended testing levels. Finally, frontloading SFR tests, performing 2 tests per eye on the same visit for the initial few visits, may be an effective strategy to increase visual field testing frequency in newly diagnosed patients, with no observable decline in performance from test fatigue.

6. Key takeaway points

SFR may not be interchangeable with SS in eyes with more severe field loss because of potentially larger differences in output sensitivity or probability score results.

FP is most likely to have the greatest impact on VF reliability followed by FN, while FL has minimal impact on reliability.

The recommended 15 % false positive cut off may be inappropriate as a threshold for judging test reliability and data inclusion in SFR. While further investigation is required, the currently recommended 15 % false positive cut off may need to be increased by as much as a factor of two when using SFR. SFR can result in a higher number of tests with poor reliability compared to SS due to seeding point errors – an artefact unique to the algorithm due to its thresholding properties.

There is minimal association between standard reliability metrics and gaze tracker metrics.

The 24–2C may be useful in flagging the presence of a clustered central visual field defect, while the 10–2 can be used to more comprehensively characterise a central field defect in some patients.

Frontloading SFR (performing 2 VF tests per eye per visit) may provide increased VF data for improved detection of progression, with the same test durations of a single SS test.

7. Conclusion

SFR, available on the HFA, provides time-saving advantages over older SITA strategies that can be leveraged to improve testing frequency; however, differences in sensitivity outputs may exist when switching between thresholding strategies, especially in patients with greater VF loss. The decrease in test duration also needs to be balanced with an increase in rates of elevated indices of unreliability, such as FP rate and artifacts. Some patients find testing conditions on SFR difficult and may ultimately be better suited for testing using SS.

8. Methods of literature search

A search of the peer-reviewed literature was conducted in February 2024 in the PubMed database (www.pubmed.org). Search terms were as follows: visual field OR perimetry AND SITA OR Swedish Interactive Threshold Algorithm. Only full-text and English language articles were included. The abstracts of 599 records underwent a title and abstract screen, of which relevant articles were selected for full-text review. This was intended to be a synthesis of the literature on the subject rather than a systematic review.

CRediT authorship contribution statement

David P. Crabb: Writing – review & editing, Visualization, Validation, Supervision, Resources, Methodology, Investigation, Formal analysis, Data curation. **Jonathan Crowston:** Writing – review & editing, Validation, Supervision, Resources, Methodology, Formal analysis, Data

curation. **Jack Phu:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Resources, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Jeremy C.K. Tan:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Resources, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Jithin Yohannan:** Writing – review & editing, Visualization, Validation, Supervision, Methodology, Formal analysis, Data curation. **Pradeep Y. Ramulu:** Writing – review & editing, Visualization, Validation, Supervision, Methodology, Investigation, Formal analysis, Data curation. **Michael Kalloniatis:** Writing – review & editing, Validation, Supervision, Resources, Methodology, Investigation, Formal analysis, Data curation, Conceptualization.

Declaration of Competing Interest

None.

References

1. Monsalve B, Ferreras A, Calvo P, et al. Diagnostic ability of Humphrey perimetry, Octopus perimetry, and optical coherence tomography for glaucomatous optic neuropathy. *Eye (Lond)*. 2017;31:443–451.
2. Heijl A, Patella VM, Chong LX, et al. A New SITA Perimetric Threshold Testing Algorithm: Construction and a Multicenter Clinical Study. *Am J Ophthalmol*. 2019; 198:154–165.
3. Phu J, Kalloniatis M. Ability of 24–2 C and 24-2 Grids to Identify Central Visual Field Defects and Structure-Function Concordance in Glaucoma and Suspects. *Am J Ophthalmol*. 2020;219:317–331.
4. Pham AT, Ramulu PY, Boland MV, Yohannan J. The Effect of Transitioning from SITA Standard to SITA Faster on Visual Field Performance. *Ophthalmology*. 2021; 128:1417–1425.
5. Phu J, Khuu SK, Agar A, Kalloniatis M. Clinical Evaluation of Swedish Interactive Thresholding Algorithm-Faster Compared With Swedish Interactive Thresholding Algorithm-Standard in Normal Subjects, Glaucoma Suspects, and Patients With Glaucoma. *Am J Ophthalmol*. 2019;208:251–264.
6. Bengtsson B, Olsson J, Heijl A, Rootzén H. A new generation of algorithms for computerized threshold perimetry, SITA. *Acta Ophthalmol Scand*. 1997;75: 368–375.
7. Bengtsson B, Heijl A. Evaluation of a new perimetric threshold strategy, SITA, in patients with manifest and suspect glaucoma. *Acta Ophthalmol Scand*. 1998;76: 268–272.
8. Bengtsson B, Heijl A. SITA Fast, a new rapid perimetric threshold test. Description of methods and evaluation in patients with manifest and suspect glaucoma. *Acta Ophthalmol Scand*. 1998;76:431–437.
9. Bengtsson B, Heijl A, Olsson J. Evaluation of a new threshold visual field strategy, SITA, in normal subjects. Swedish Interactive Thresholding Algorithm. *Acta Ophthalmol Scand*. 1998;76:165–169.
10. Budenz DL, Rhee P, Feuer WJ, McSoley J, Johnson CA, Anderson DR. Sensitivity and specificity of the Swedish interactive threshold algorithm for glaucomatous visual field defects. *Ophthalmology*. 2002;109:1052–1058.
11. Sekhar GC, Naduvilath TJ, Lakkai M, et al. Sensitivity of Swedish interactive threshold algorithm compared with standard full threshold algorithm in Humphrey visual field testing. *Ophthalmology*. 2000;107:1303–1308.
12. Sharma AK, Goldberg I, Graham SL, Mohsin M. Comparison of the Humphrey swedish interactive thresholding algorithm (SITA) and full threshold strategies. *J Glaucoma*. 2000;9:20–27.
13. Shirato S, Inoue R, Fukushima K, Suzuki Y. Clinical evaluation of SITA: a new family of perimetric testing strategies. *Graefes Arch Clin Exp Ophthalmol*. 1999;237: 29–34.
14. Wild JM, Pacey IE, O'Neill EC, Cunliffe IA. The SITA perimetric threshold algorithms in glaucoma. *Invest Ophthalmol Vis Sci*. 1999;40:1998–2009.
15. Bengtsson B, Heijl A. Comparing significance and magnitude of glaucomatous visual field defects using the SITA and Full Threshold strategies. *Acta Ophthalmol Scand*. 1999;77:143–146.
16. Aoki Y, Takahashi G, Kitahara K. Comparison of Swedish interactive threshold algorithm and full threshold algorithm for glaucomatous visual field loss. *Eur J Ophthalmol*. 2007;17:196–202.
17. Budenz DL, Rhee P, Feuer WJ, McSoley J, Johnson CA, Anderson DR. Comparison of glaucomatous visual field defects using standard full threshold and Swedish interactive threshold algorithms. *Arch Ophthalmol*. 2002;120:1136–1141.
18. Hirasawa K, Shoji N. Swedish Interactive Threshold Algorithm for central visual field defects unrelated to nerve fiber layer. *Graefes Arch Clin Exp Ophthalmol*. 2016; 254:845–854.
19. Bourne RR, Jahanbakhsh K, Boden C, et al. Reproducibility of visual field end point criteria for standard automated perimetry, full-threshold, and Swedish interactive thresholding algorithm strategies: diagnostic innovations in glaucoma study. *Am J Ophthalmol*. 2007;144:908–913.
20. Musch DC, Gillespie BW, Motyka BM, Niziol LM, Mills RP, Lichter PR. Converting to SITA-standard from full-threshold visual field testing in the follow-up phase of a clinical trial. *Invest Ophthalmol Vis Sci*. 2005;46:2755–2759.
21. Schimiti RB, Avelino RR, Kara-José N, Costa VP. Full-threshold versus Swedish Interactive Threshold Algorithm (SITA) in normal individuals undergoing automated perimetry for the first time. *Ophthalmology*. 2002;109:2084–2092. discussion 92.
22. Bengtsson B, Heijl A. Inter-subject variability and normal limits of the SITA Standard, SITA Fast, and the Humphrey Full Threshold computerized perimetry strategies, SITA STATPAC. *Acta Ophthalmol Scand*. 1999;77:125–129.
23. Johnson CA, Sherman K, Doyle C, Wall M. A comparison of false-negative responses for full threshold and SITA standard perimetry in glaucoma patients and normal observers. *J Glaucoma*. 2014;23:288–292.
24. Heijl A, Patella VM, Flanagan JG, et al. False Positive Responses in Standard Automated Perimetry. *Am J Ophthalmol*. 2022;233:180–188.
25. Costa VP, Zangalli CS, Jammal AA, et al. 24-2 SITA Standard versus 24-2 SITA Faster in Perimetry-Naive Normal Subjects. *Ophthalmol Glaucoma*. 2023;6: 129–136.
26. Lavanya R, Riyazuddin M, Dasari S, et al. A Comparison of the Visual Field Parameters of SITA Faster and SITA Standard Strategies in Glaucoma. *J Glaucoma*. 2020;29:783–788.
27. Mendieta N, Suárez J, Blasco C, Muñiz R, Pueyo C. A Comparative Study between Swedish Interactive Thresholding Algorithm Faster and Swedish Interactive Thresholding Algorithm Standard in Glaucoma Patients. *J Curr Ophthalmol*. 2021; 33:247–252.
28. Thulasidas M, Patyal S. Comparison of 24-2 Faster, Fast, and Standard Programs of Swedish Interactive Threshold Algorithm of Humphrey Field Analyzer for Perimetry in Patients With Manifest and Suspect Glaucoma. *J Glaucoma*. 2020;29: 1070–1076.
29. Rodríguez-Agirretxe I, Loizate E, Astorkiza B, Onaindia A, Galdos-Olasagasti L, Basasoro A. Validation of the SITA faster strategy for the management of glaucoma. *Int Ophthalmol*. 2022;42:2347–2354.
30. Heijl A. Time changes of contrast thresholds during automatic perimetry. *Acta Ophthalmol (Copenh)*. 1977;55:696–708.
31. Heijl A, Drance SM. Changes in differential threshold in patients with glaucoma during prolonged perimetry. *Br J Ophthalmol*. 1983;67:512–516.
32. Le CT, Fiksel J, Ramulu P, Yohannan J. Differences in visual field loss pattern when transitioning from SITA standard to SITA faster. *Sci Rep*. 2022;12:7001.
33. Qian CX, Chen Q, Cun Q, et al. Comparison of the SITA Faster-a new visual field strategy with SITA Fast strategy. *Int J Ophthalmol*. 2021;14:1185–1191.
34. Tan JCK, Phu J, Go D, et al. Evaluation of the consistency of glaucomatous visual field defects using a clustered SITA-Faster protocol. *Ophthalmology*. 2023.
35. Bengtsson B. Reliability of computerized perimetric threshold tests as assessed by reliability indices and threshold reproducibility in patients with suspect and manifest glaucoma. *Acta Ophthalmol Scand*. 2000;78:519–522.
36. Camp AS, Long CP, Patella VM, Proudfoot JA, Weinreb RN. Standard Reliability and Gaze Tracking Metrics in Glaucoma and Glaucoma Suspects. *Am J Ophthalmol*. 2022;234:91–98.
37. Phu J, Kalloniatis M. The Frontloading Fields Study: The Impact of False Positives and Seeding Point Errors on Visual Field Reliability When Using SITA-Faster. *Transl Vis Sci Technol*. 2022;11:20.
38. Phu J, Kalloniatis M. Gaze tracker parameters have little association with visual field metrics of intrasession frontloaded SITA-Faster 24-2 visual field results. *Ophthalmic Physiol Opt*. 2022;42:973–985.
39. Rabiolo A, Morales E, Afifi AA, Yu F, Nouri-Mahdavi K, Caprioli J. Quantification of Visual Field Variability in Glaucoma: Implications for Visual Field Prediction and Modeling. *Transl Vis Sci Technol*. 2019;8:25.
40. Stagg B, Mariottoni EB, Berchuck S, et al. Longitudinal visual field variability and the ability to detect glaucoma progression in black and white individuals. *Br J Ophthalmol 106:1115-20*. 2022.
41. Heijl A, Lindgren A, Lindgren G. Test-retest variability in glaucomatous visual fields. *Am J Ophthalmol*. 1989;108:130–135.
42. Russell RA, Crabb DP, Malik R, Garway-Heath DF. The relationship between variability and sensitivity in large-scale longitudinal visual field data. *Invest Ophthalmol Vis Sci*. 2012;53:5985–5990.
43. Tan JCK, Agar A, Kalloniatis M, Phu J. Quantification and Predictors of Visual Field Variability in Healthy, Glaucoma Suspect, and Glaucomatous Eyes Using SITA-Faster. *Ophthalmology*. 2024;131:658–666.
44. Lindgren A. *Regression analysis of censored data with applications in perimetry* (Thesis). Lund University; 1999.
45. Artes PH, Iwase A, Ohno Y, Kitazawa Y, Chauhan BC. Properties of perimetric threshold estimates from Full Threshold, SITA Standard, and SITA Fast strategies. *Invest Ophthalmol Vis Sci*. 2002;43:2654–2659.
46. Saunders LJ, Russell RA, Crabb DP. Measurement precision in a series of visual fields acquired by the standard and fast versions of the Swedish interactive thresholding algorithm: analysis of large-scale data from clinics. *JAMA Ophthalmol*. 2015;133:74–80.
47. Giammaria S, Vianna JR, Ohno Y, Iwase A, Chauhan BC. Nomograms for Converting Perimetric Sensitivity From Full Threshold and SITA Fast to SITA Standard in Patients With Glaucoma and Healthy Subjects. *Transl Vis Sci Technol*. 2021;10:2.
48. Matsuura M, Hirasawa K, Murata H, Asaoka R. The Relationship Between Visual Acuity and the Reproducibility of Visual Field Measurements in Glaucoma Patients. *Invest Ophthalmol Vis Sci*. 2015;56:5630–5635.
49. Kim JH, Rabiolo A, Morales E, et al. Risk Factors for Fast Visual Field Progression in Glaucoma. *Am J Ophthalmol*. 2019;207:268–278.

50. Wall M, Doyle CK, Zamba KD, Artes P, Johnson CA. The repeatability of mean defect with size III and size V standard automated perimetry. *Invest Ophthalmol Vis Sci*. 2013;54:1345–1351.
51. Wall M, Woodward KR, Doyle CK, Artes PH. Repeatability of automated perimetry: a comparison between standard automated perimetry with stimulus size III and V, matrix, and motion perimetry. *Invest Ophthalmol Vis Sci*. 2009;50:974–979.
52. Junoy Montolio FG, Wesselink C, Gordijn M, Jansonius NM. Factors that influence standard automated perimetry test results in glaucoma: test reliability, technician experience, time of day, and season. *Invest Ophthalmol Vis Sci*. 2012;53:7010–7017.
53. WuDunn D, Takusagawa HL, Rosdahl JA, et al. Central Visual Field Testing in Early Glaucoma: A Report by the American Academy of. *Ophthalmol Ophthalmol*. 2024;131:240–248.
54. De Moraes CG, Hood DC, Thenappan A, et al. 24-2 Visual Fields Miss Central Defects Shown on 10-2 Tests in Glaucoma Suspects, Ocular Hypertensives, and Early Glaucoma. *Ophthalmology*. 2017;124:1449–1456.
55. West ME, Sharpe GP, Hutchison DM, et al. Value of 10-2 Visual Field Testing in Glaucoma Patients with Early 24-2 Visual Field Loss. *Ophthalmology*. 2021;128:545–553.
56. Park SC, Kung Y, Su D, et al. Parafoveal scotoma progression in glaucoma: humphrey 10-2 versus 24-2 visual field analysis. *Ophthalmology*. 2013;120:1546–1550.
57. Phu J, Rafla D, Kalloniatis M. Which glaucoma patients benefit from 10-2 visual field testing? Proposing the functional vulnerability zone framework. *Clin Exp Optom:1-13*. 2023.
58. Fung SS, Lemer C, Russell RA, Malik R, Crabb DP. Are practical recommendations practiced? A national multi-centre cross-sectional study on frequency of visual field testing in glaucoma. *Br J Ophthalmol*. 2013;97:843–847.
59. Chen S, McKendrick AM, Turpin A. Choosing two points to add to the 24-2 pattern to better describe macular visual field damage due to glaucoma. *Br J Ophthalmol*. 2015;99:1236–1239.
60. Ehrlich AC, Raza AS, Ritch R, Hood DC. Modifying the Conventional Visual Field Test Pattern to Improve the Detection of Early Glaucomatous Defects in the Central 10°. *Transl Vis Sci Technol*. 2014;3(6).
61. Rafla D, Kalloniatis M, Phu J. The effect of macular visual field test density on central structure-function concordance in glaucoma. *Clin Exp Optom:1-10*. 2024.
62. Behera G, Nath A, Ramasamy A, Kaliaperumal S. Comparing Static Perimetry Protocols of Central Field Testing among Patients with Glaucoma. *Optom Vis Sci*. 2023;100:406–411.
63. Hong JW, Baek MS, Lee JY, Song MK, Shin JW, Kook MS. Comparison of the 24-2 and 24–2 C Visual Field Grids in Determining the Macular Structure-Function Relationship in Glaucoma. *J Glaucoma*. 2021;30:887–894.
64. Phu J, Kalloniatis M. Comparison of 10-2 and 24–2 C Test Grids for Identifying Central Visual Field Defects in Glaucoma and Suspect Patients. *Ophthalmology*. 2021;128:1405–1416.
65. Bengtsson B, Heijl A. False-negative responses in glaucoma perimetry: indicators of patient performance or test reliability? *Invest Ophthalmol Vis Sci*. 2000;41:2201–2204.
66. Katz J, Sommer A. Reliability indexes of automated perimetric tests. *Arch Ophthalmol*. 1988;106:1252–1254.
67. Rao HL, Yadav RK, Begum VU, et al. Role of visual field reliability indices in ruling out glaucoma. *JAMA Ophthalmol*. 2015;133:40–44.
68. Olsson J, Bengtsson B, Heijl A, Rootzén H. An improved method to estimate frequency of false positive answers in computerized perimetry. *Acta Ophthalmol Scand*. 1997;75:181–183.
69. Garway-Heath DF, Quartilho A, Prah P, Crabb DP, Cheng Q, Zhu H. Evaluation of Visual Field and Imaging Outcomes for Glaucoma Clinical Trials (An American Ophthalomological Society Thesis). *Trans Am Ophthalmol Soc*. 2017;115:T4.
70. Johnson CA, Keltner JL, Cello KE, et al. Baseline visual field characteristics in the ocular hypertension treatment study. *Ophthalmology*. 2002;109:432–437.
71. Aboobakar IF, Wang J, Chauhan BC, et al. Factors Predicting a Greater Likelihood of Poor Visual Field Reliability in Glaucoma Patients and Suspects. *Transl Vis Sci Technol*. 2020;9:4.
72. Yohannan J, Wang J, Brown J, et al. Evidence-based Criteria for Assessment of Visual Field Reliability. *Ophthalmology*. 2017;124:1612–1620.
73. Ishiyama Y, Murata H, Asaoka R. The Usefulness of Gaze Tracking as an Index of Visual Field Reliability in Glaucoma Patients. *Invest Ophthalmol Vis Sci*. 2015;56:6233–6236.
74. Asaoka R, Fujino Y, Aoki S, Matsuura M, Murata H. Estimating the Reliability of Glaucomatous Visual Field for the Accurate Assessment of Progression Using the Gaze-Tracking and Reliability Indices. *Ophthalmol Glaucoma*. 2019;2:111–119.
75. Ishiyama Y, Murata H, Mayama C, Asaoka R. An objective evaluation of gaze tracking in Humphrey perimetry and the relation with the reproducibility of visual fields: a pilot study in glaucoma. *Invest Ophthalmol Vis Sci*. 2014;55:8149–8152.
76. Phu J, Kalloniatis M. A Strategy for Seeding Point Error Assessment for Retesting (SPEAR) in Perimetry Applied to Normal Subjects, Glaucoma Suspects, and Patients With Glaucoma. *Am J Ophthalmol*. 2021;221:115–130.
77. Gardiner SK, Demirel S, Johnson CA. Is there evidence for continued learning over multiple years in perimetry? *Optom Vis Sci*. 2008;85:1043–1048.
78. Rubinstein NJ, Turpin A, Denniss J, McKendrick AM. Effects of Criterion Bias on Perimetric Sensitivity and Response Variability in Glaucoma. *Transl Vis Sci Technol*. 2021;10:18.
79. Gardiner SK, Swanson WH, Mansberger SL. Long- and Short-Term Variability of Perimetry in Glaucoma. *Transl Vis Sci Technol*. 2022;11:3.
80. Heijl A, Bengtsson B. The effect of perimetric experience in patients with glaucoma. *Arch Ophthalmol*. 1996;114:19–22.
81. Heijl A, Lindgren G, Olsson J. The effect of perimetric experience in normal subjects. *Arch Ophthalmol*. 1989;107:81–86.
82. Wild JM, Searle AE, Dengler-Harles M, O'Neill EC. Long-term follow-up of baseline learning and fatigue effects in the automated perimetry of glaucoma and ocular hypertensive patients. *Acta Ophthalmol (Copenh)*. 1991;69:210–216.
83. Castro DP, Kawase J, Melo Jr LA. Learning effect of standard automated perimetry in healthy individuals. *Arq Bras Oftalmol*. 2008;71:523–528.
84. Barkana Y, Gerber Y, Mora R, Liebmann JM, Ritch R. Effect of eye testing order on automated perimetry results using the Swedish Interactive Threshold Algorithm standard 24-2. *Arch Ophthalmol*. 2006;124:781–784.
85. Kelly SR, Bryan SR, Crabb DP. Does eye examination order for standard automated perimetry matter? *Acta Ophthalmol*. 2019;97:e833–e838.
86. Tan JCK, Kalloniatis M, Phu J. Frontloading SITA-Faster Can Increase Frequency and Reliability of Visual Field Testing at Minimal Time Cost. *Ophthalmol Glaucoma*. 2023;6:445–456.
87. Hutchings N, Wild JM, Hussey MK, Flanagan JG, Trope GE. The long-term fluctuation of the visual field in stable glaucoma. *Invest Ophthalmol Vis Sci*. 2000;41:3429–3436.
88. Rabiolo A, Morales E, Kim JH, et al. Predictors of Long-Term Visual Field Fluctuation in Glaucoma Patients. *Ophthalmology*. 2020;127:739–747.
89. Crabb DP, Garway-Heath DF. Intervals between visual field tests when monitoring the glaucomatous patient: wait-and-see approach. *Invest Ophthalmol Vis Sci*. 2012;53(6):2770.
90. Chauhan BC, Garway-Heath DF, Goñi FJ, et al. Practical recommendations for measuring rates of visual field change in glaucoma. *Br J Ophthalmol*. 2008;92:569–573.
91. 2017. European Glaucoma Society Terminology and Guidelines for Glaucoma, 4th Edition - Chapter 3: Treatment principles and options Supported by the EGS Foundation: Part 1: Foreword; Introduction; Glossary; Chapter 3 Treatment principles and options. *Br J Ophthalmol* 101:130-195.
92. Crabb D.P., Russell R.A., Malik R., Anand N., Baker H., et al. 2014. Health Services and Delivery Research. In Frequency of visual field testing when monitoring patients newly diagnosed with glaucoma: mixed methods and modelling. Southampton (UK): NIHR Journals Library Copyright © Queen's Printer and Controller of HMSO 2014. This work was produced by Crabb et al. under the terms of a commissioning contract issued by the Secretary of State for Health. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK. Number of.
93. Stagg BC, Stein JD, Medeiros FA, et al. The Frequency of Visual Field Testing in a US Nationwide Cohort of Individuals with Open-Angle Glaucoma. *Ophthalmol Glaucoma*. 2022;5:587–593.
94. Gedde SJ, Vinod K, Wright MM, et al. Primary Open-Angle Glaucoma Preferred Practice Pattern®. *Ophthalmology*. 2021;128. P71-p150.
95. Herbert P, Hou K, Bradley C, et al. Forecasting Risk of Future Rapid Glaucoma Worsening Using Early Visual Field, OCT, and Clinical Data. *Ophthalmol Glaucoma*. 2023;6:466–473.
96. Shuldiner SR, Boland MV, Ramulu PY, et al. Predicting eyes at risk for rapid glaucoma progression based on an initial visual field test using machine learning. *PLoS One*. 2021;16, e0249856.
97. Bradley C, Herbert P, Hou K, Unberath M, Ramulu P, Yohannan J. Comparing the Accuracy of Peripapillary OCT Scans and Visual Fields to Detect Glaucoma Worsening. *Ophthalmology*. 2023;130:631–639.
98. Bradley C, Hou K, Herbert P, et al. Evidence-Based Guidelines for the Number of Peripapillary OCT Scans Needed to Detect Glaucoma Worsening. *Ophthalmology*. 2023;130:39–47.
99. Garway-Heath DF, Crabb DP, Bunce C, et al. Latanoprost for open-angle glaucoma (UKGTS): a randomised, multicentre, placebo-controlled trial. *Lancet*. 2015;385:1295–1304.
100. Garway-Heath DF, Lascaratos G, Bunce C, Crabb DP, Russell RA, Shah A. The United Kingdom Glaucoma Treatment Study: a multicenter, randomized, placebo-controlled clinical trial: design and methodology. *Ophthalmology*. 2013;120:68–76.
101. Phu J, Kalloniatis M. Viability of Performing Multiple 24-2 Visual Field Examinations at the Same Clinical Visit: The Frontloading Fields Study (FFS). *Am J Ophthalmol*. 2021;230:48–59.
102. Wang H, Kalloniatis M, Tan JCK, Phu J. Frontloading visual field tests detect earlier mean deviation progression when applied to real-world-derived early-stage glaucoma data. *Ophthalmic Physiol Opt*. 2024;44:426–441.
103. Phu J, Kalloniatis M. The Frontloading Fields Study (FFS): Detecting Changes in Mean Deviation in Glaucoma Using Multiple Visual Field Tests Per Clinical Visit. *Transl Vis Sci Technol*. 2021;10:21.
104. Phu J, Tan J, Kalloniatis M. Multiple (frontloaded) visual field tests increase identification of very slow mean deviation progression in glaucoma. *Can J Ophthalmol*. 2023.
105. Heijl A, Buchholz P, Norrgren G, Bengtsson B. Rates of visual field progression in clinical glaucoma care. *Acta Ophthalmol*. 2013;91:406–412.
106. Chauhan BC, Malik R, Shuba LM, Rafuse PE, Nicolela MT, Artes PH. Rates of glaucomatous visual field change in a large clinical population. *Invest Ophthalmol Vis Sci*. 2014;55:4135–4143.
107. Tan JCK, Phu J, Go D, et al. Evaluation of the Consistency of Glaucomatous Visual Field Defects Using a Clustered SITA-Faster Protocol. *Ophthalmology*. 2023;130:1138–1148.

108. Phu J, Kalloniatis M. Patient and technician perspectives following the introduction of frontloaded visual field testing in glaucoma assessment. *Clin Exp Optom*. 2022;105:617–623.
109. Jones PR, Campbell P, Callaghan T, et al. Glaucoma Home Monitoring Using a Tablet-Based Visual Field Test (Eyecatcher): An Assessment of Accuracy and Adherence Over 6 Months. *Am J Ophthalmol*. 2021;223:42–52.
110. Kang J, De Arrigunaga S, Freeman SE, et al. Comparison of Perimetric Outcomes from a Tablet Perimeter, Smart Visual Function Analyzer, and Humphrey Field Analyzer. *Ophthalmol Glaucoma*. 2023;6:509–520.
111. Phu J, Wang H, Kalloniatis M. Comparing a head-mounted virtual reality perimeter and the Humphrey Field Analyzer for visual field testing in healthy and glaucoma patients. *Ophthalmic Physiol Opt*. 2024;44:83–95.
112. Prea SM, Kong GYX, Guymer RH, Vingrys AJ. Uptake, Persistence, and Performance of Weekly Home Monitoring of Visual Field in a Large Cohort of Patients With Glaucoma. *Am J Ophthalmol*. 2021;223:286–295.
113. Ichhpujani P, Dhillon H. Spotlight on iPad Visual Field Tests Efficacy. *Clin Ophthalmol*. 2022;16:2179–2185.
114. Selvan K, Mina M, Abdelmeguid H, Gulsha M, Vincent A, Sarhan A. Virtual reality headsets for perimetry testing: a systematic review. *Eye (Lond)*. 2024;38:1041–1064.