A novel Bayesian adaptive method for mapping the visual field

Pengjing Xu

College of Optometry, The Ohio State University, Columbus, OH, USA ✉

Luis Andres Lesmes

Adaptive Sensory Technology, Inc., San Diego, CA, USA ✉

Deyue Yu

College of Optometry, The Ohio State University, Columbus, OH, USA ✉

Zhong-Lin Lu

Division of Arts and Sciences, NYU Shanghai, Shanghai, China
Center for Neural Science and Department of Psychology, New York University, New York, NY, USA
NYU-ECNU Institute of Cognitive Neuroscience at NYU Shanghai, Shanghai, China ✉

Measuring visual functions such as light and contrast sensitivity, visual acuity, reading speed, and crowding across retinal locations provides visual-field maps (VFMs) that are extremely valuable for detecting and managing eye diseases. Although mapping light sensitivity is a standard glaucoma test, the measurement is often noisy (Keltner et al., 2000). Mapping other visual functions is even more challenging. To improve the precision of light-sensitivity mapping and enable other VFM assessments, we developed a novel hybrid Bayesian adaptive testing framework, the qVFM method. The method combines a global module for preliminary assessment of the VFM's shape and a local module for assessing individual visual-field locations. This study validates the qVFM method in measuring light sensitivity across the visual field. In both simulation and psychophysics studies, we sampled 100 visual-field locations ($60° \times 60°$) and compared the performance of qVFM with the qYN procedure (Lesmes et al., 2015) that measured light sensitivity at each location independently. In the simulations, a simulated observer was tested monocularly for 1,000 runs with 1,200 trials/run, to compare the accuracy and precision of the two methods. In the experiments, data were collected from 12 eyes (six left, six right) of six human subjects. Subjects were cued to report the presence or absence of a target stimulus, with the luminance and location of the target adaptively selected in each trial. Both simulations and a psychological experiment showed that the qVFM method can provide accurate, precise, and efficient mapping of light sensitivity. This method can be extended to map other visual functions, with potential clinical signals for monitoring vision loss, evaluating therapeutic interventions, and developing effective rehabilitation for low vision.

Introduction

Peripheral vision is important for performing a wide range of daily activities (Higgitt & Smith, 1955; Larson & Loschky, 2009; Lemmink, Dijkstra, & Visscher, 2005; Strasburger, Rentschler, & Jüttner, 2011). A number of conditions can significantly affect peripheral vision, including glaucoma (Caprioli, 1991; Ng et al., 2012; Smith, Katz, & Quigley, 1996), diabetic retinopathy (Bengtsson, Heijl, & Agardh, 2005), stroke (Townend et al., 2007), pituitary disease (Okamoto, Okamoto, Hiraoka, Yamada, & Oshika, 2008; Rowe et al., 2015), brain tumors (Huber, 1976), and other neurological deficits (Papageorgiou et al., 2007). People with damaged peripheral vision often experience major difficulties in performing many important activities, such as reading (Higgitt & Smith, 1955; Ramulu, West, Munoz, Jampel, & Friedman, 2009), driving (Johnson & Keltner, 1983; Wood, 2002), spatial orienting (Marron & Bailey, 1982), mobility (Lovie-Kitchin, Soong, Hassan, & Woods, 2010; Lovie-Kitchin, Woods, Hassan, & Soong, 2001), scene-gist recognition (Larson & Loschky, 2009), social interactions (Rogers & Landers, 2005), and tool use/manipulation (National

Research Council, 2002; P. Ramulu, 2009; Wang, Javitt, Rowe, & Meng, 1996).

The visual-field map (VFM), defined as "the spatial array of visual sensations available to observation in introspectionist psychological experiments" (Smythies, 1996, pp. 369–371)—analogous to the concept of the field of view of optical instruments and sensors—provides a quantitative assessment of an individual's central and peripheral vision through measurements of visual function(s) across retinal locations (Aulhorn & Harms, 1972; Strasburger et al., 2011). The VFM is used not only to detect visual dysfunctions in the clinic but also by the Social Security Administration as part of their procedure for determining visual disability (National Research Council, 2002; Thornton et al., 2004; Wixon & Strand, 2013).

As of now, the VFM is typically measured with perimetry (Goldmann, 1945b, 1945a; Harms, 1952), which usually involves measurement of visual sensitivity in detecting a small light target projected on a uniform dark background (Aulhorn & Harms, 1972). Since peripheral vision is heavily involved in many different visual activities, a comprehensive characterization of peripheral vision requires assessment of VFMs of multiple visual functions in addition to light sensitivity (Strasburger et al., 2011). Such characterization is crucial for monitoring the status of vision loss, for developing and providing effective rehabilitation interventions, and for obtaining projections of potential benefits from interventions (Advanced Glaucoma Intervention Study Investigators, 1994).

While the majority of eye-care practitioners use standard automated perimetry (SAP) as part of their clinical ophthalmic diagnostic procedure (Anderson & Patella, 1999; Chris A. Johnson, Wall, & Thompson, 2011), a number of different types of perimetry test are also available, including frequency-doubling technology perimetry (Chauhan & Johnson, 1999; Johnson & Samuels, 1997), Amsler grid (Easterbrook, 1984; Fink & Sadun, 2004; Nguyen et al., 2009; Wall & May, 1987), color perimetry (Carlow, Flynn, & Shipley, 1976; Hart, Hartz, Hagen, & Clark, 1984; Sample & Weinreb, 1990), flicker perimetry (Lachenmayr, Drance, Douglas, & Mikelberg, 1991; Lachenmayr et al., 1994; Yoshiyama & Johnson, 1997), motion perimetry (Wall, Brito, & Kutzko, 1997; Wall & Ketoff, 1995; Wall, Woodward, Doyle, & Artes, 2009), high-pass resolution perimetry (Chauhan, House, McCormick, & LeBlanc, 1999; FrisÉn, 1993; Frisén, 1992) , multifocal visual-evoked potential (Hood, Odel, & Winn, 2003; James, 2003; Klistorner, Graham, Grigg, & Billson, 1998), pupil perimetry (Kardon, 1992; Kardon, Kirkali, & Thompson, 1991; Rajan, Bremner, & Riordan-Eva, 2002), and rarebit perimetry (Brusini, Salvetat, Parisi, & Zeppieri, 2005; Martin, 2005; Martin & Wanger, 2004). The 1994 report of the Committee on Vision

provides an overview of VFM measurement techniques and factors that are relevant to VFM testing (Lennie, 1994). Because it is difficult to balance test efficiency and precision, many of these VFM tests are still considered as screening references or research tools and are not frequently used in the clinic (Johnson et al., 2011). The current in-clinic evaluation of ophthalmic disorders mostly consists of measurements of multiple visual functions at a single visual-field location, which provides a rather limited characterization of residual spatial vision, and mapping visual functions beyond light sensitivity is uncommon (Broadway, 2012; Hodapp, Parrish, & Anderson, 1993; Lim, Mitchell, Seddon, Holz, & Wong, 2012; Markowitz, 2006).

Ideally, a comprehensive test of peripheral vision should include light sensitivity and many of the visual functions in the various types of perimetry tests just discussed, as well as contrast sensitivity (Daitch & Green, 1969; Swanson, Malinovsky, et al., 2014), visual acuity (Thompson, Montague, Cox, & Corbett, 1982; VA, 1965), color vision (Carlow et al., 1976; Hart et al., 1984; Sample & Weinreb, 1990, 1992), reading speed (Ramulu et al., 2009; Yu, Cheung, Legge, & Chung, 2010), crowding (Balas, Nakano, & Rosenholtz, 2009; Levi & Carney, 2009), and others. However, there are many challenges. A precise and accurate VFM assessment with existing methods is extremely time consuming (Artes, Iwase, Ohno, Kitazawa, & Chauhan, 2002; Weinreb & Kaufman, 2009, 2011). In fact, the results from SAP are noisy (Keltner et al., 2000; Stewart & Hunt, 1993); assessment of the VFM of other visual functions is difficult and rarely used in clinic. Although many new perimetric methods can provide helpful clinical information, they have not been sufficiently validated for routine clinical use (Johnson et al., 2011; Keltgen & Swanson, 2012; Strasburger et al., 2011; Swanson, Malinovsky, et al., 2014).

Currently, SAP is the gold standard for detecting optic-nerve disease in glaucoma and staging glaucomatous damage into different categories (Broadway, 2012; Caprioli, 1991; Drance & Anderson, 1985; Ng et al., 2012; Sharma, Sample, Zangwill, & Schuman, 2008). Staging visual-field damage is crucial for optimal treatment and for establishing the rate and risk of progression of each subtype of glaucomatous visual-field loss ("ICD-10 Glaucoma Staging Definitions," 2015; Katz, Gilbert, Quigley, & Sommer, 1997; Mills et al., 2006; Quigley, Tielsch, Katz, & Sommer, 1996; Susanna & Vessani, 2009). However, staging based on perimetry is difficult because of the inaccuracy, inefficiency, and lack of robustness of existing SAP devices (Heijl, Lindgren, & Lindgren, 1988, 1989; Keltner et al., 2000). The variability of the estimated visual sensitivity from SAP is relatively large for individuals with healthy vision. Heijl, Lindgren, and Olsson (1987) showed that the intratest variation

(short-term fluctuation) was 1.59 dB and intertest variation (standard deviation) was about 4 dB in the peripheral visual field. Katz and Sommer (1987) found that the standard deviation ranged from 1.2 to 2.8 dB and from 2.2 to 8.3 dB for subjects under and over 60 years of age, respectively, and short-term fluctuations contributed to 25% of the variability of threshold measurements over time. Lewis, Johnson, Keltner, and Labermeier (1986) measured the test–retest variability of the estimated visual sensitivity in different regions of the visual field of healthy subjects using six commercially available automated threshold static perimeters, and found that the standard deviation ranged from 1.23 to 2.25 dB. Werner, Saheb, and Thomas (1982) found that the estimated light-detection threshold had a test–retest variability of more than 4 dB in at least one retinal location in 11 of 12 healthy eyes. Wilensky and Joondeph (1984) concluded that estimated visual sensitivity from automated visual-field tests must be interpreted cautiously, because an apparent pathologic change may represent only normal variation.

It has been reported that 85.9% of the abnormal visual fields determined by SAP in the Ocular Hypertension Treatment Study were not verified on retest (Keltner et al., 2000). For individuals with progressive glaucoma, the uncertainty exhibited by SAP makes it difficult to distinguish between truly progressive glaucoma and long-term variability unless several VFMs are obtained over time (Ashimatey & Swanson, 2016; Swanson, Horner, Dul, & Malinovsky, 2014; Weinreb & Kaufman, 2009). Werner et al. (1982) concluded that a minimum of six repeated VFM measurements were required to make informed clinical judgments as to whether a patient's VFM was stable or progressing. Quantitative approaches using linear regression have come to similar conclusions, indicating that approximately seven repeated VFM measurements obtained over several years are needed to reliably distinguish progression from intratest variability (Birch, Wishart, & O'Donnell, 1995; Katz et al., 1997; Smith et al., 1996). For studies using a discrete measure (change from baseline) rather than regression techniques, it has been found that confirmation of change is necessary to avoid "overcalling" progression of visual-field loss. In the Normal-Tension Glaucoma Study, Schulzer et al. (1994) found that four to six confirming VFM tests (two of three tests performed within 1 to 4 weeks showing change, followed by two of three tests performed 3 months later) were required to reliably determine visual-field progression. Because of the variabilities in SAP, only large changes of light sensitivity can be used as a clinically relevant functional endpoint by the FDA. For example, a 24-2 (Humphrey field analyzer) full VFM of light sensitivity is considered to demonstrate progression if five or more reproducible points of the 52 non-blind-spot visual-field locations have significant changes from the baseline in the glaucoma-change probability analysis (Lloyd, Harris, Wadhwa, & Chambers, 2008; Lloyd, Wadhwa, Eydelman, & Kramm, 2011). Alternatively, visual-field progression is suspected if the mean difference in threshold for the entire field is statistically and clinically significant between treatment and no-treatment groups (7-dB change on more than one examination; Weinreb & Kaufman, 2009, 2011).

SAP often involves adaptive psychophysical procedures that use observer responses to focus stimulus presentation to predefined regions of the empirical psychometric function (Anderson & Patella, 1999; Leek, 2001; Treutwein, 1995). Most SAP devices implement the Full Threshold algorithm (Haley, 1986; Heijl, 1985), which is a nonparametric 1-up/1-down staircase method for estimating independent thresholds around 50% correct (or "yes") at predetermined test locations (Dixon & Mood, 1948; Anderson & Patella, 1999). The step size(s), number of reversals and number of staircase procedures at each test location influence the accuracy of the estimated thresholds. To achieve an acceptable accuracy level, tests using the Full Threshold algorithm are very time consuming. Test durations up to 20 minutes are not unusual for patients with glaucomatous visual-field defects (Haley, 1986; Heijl, 1985). Shorter tests are available, but generally with reduced accuracy (Heijl, 1977; Johnson, Chauhan, & Shapiro, 1992). For example, one shorter test algorithm, FASTPAC, has been reported to produce less reliable results than the Full Threshold algorithm (Flanagan, Wild, & Trope, 1993). Another shorter test algorithm, SITA (the Swedish Interactive Thresholding Algorithm), can produce the same quality of test results as the Full Threshold strategy with considerable reduction of test time. However, it can only be used with the Goldmann size III stimulus of the Humphrey perimeter, and was released only for glaucomatous patients because its a priori threshold distribution was based on glaucoma (Artes et al., 2002; Bengtsson & Heijl, 1998). More features of different threshold test algorithms are detailed in the Discussion section.

To improve the accuracy and precision of light-sensitivity VFM, enable other VFM testing, and reduce testing time, we have developed a novel Bayesian adaptive testing framework, the qVFM method, that combines a global module for preliminary assessment of the general shape of the VFM, a local module for assessing visual functions at each individual visual-field location, and a switch module that evaluates the rate of information gain in the global module and determines when to switch to the local module (Lu, Xu, Lesmes, & Yu, n.d.). Taking advantage of a one-step-ahead search procedure to gain maximum information on the VFM, the qVFM method selects the optimal stimulus location and intensity to test in each trial.

The Bayesian framework was first applied in the landmark development of the QUEST method and is now extensively used in psychophysics (Kujala & Lukka, 2006; Leek, 2001; Lu & Dosher, 2013; Treutwein, 1995; Watson, 2017). Whereas QUEST was designed to measure a single threshold (King-Smith, Grigsby, Vingrys, Benes, & Supowit, 1994; Simpson, 1989; Watson & Pelli, 1983), recent developments have extended the adaptive procedure to measure various psychological functions, including the threshold-versus-contrast external noise function (Lesmes, Jeon, Lu, & Dosher, 2006), contrast sensitivity function (Dorr et al., 2015; Hou et al., 2010; Hou, Lesmes, Bex, Dorr, & Lu, 2015; Lesmes, Lu, Baek, & Albright, 2010), sensory memory decay (Baek, Lesmes, & Lu, 2014, 2016), reading function (Hou et al., 2018; Shepard et al., 2019), and detailed time course of perceptual sensitivity change (Lu, Zhang, Zhao, & Dosher, 2018; Zhao, Lesmes, & Lu, 2019).

To achieve better efficiency in measuring detection thresholds and be consistent with SAP, we adapted a yes/no (YN) task in this study. The primary statistical advantage of the YN task is the wider dynamic range of the YN psychometric function (Klein, 2001; Leek, 2001; Lesmes et al., 2015). The notable disadvantage of the YN task is the significant contribution of the observer's decision criterion (response bias) in detection behavior. Previously, we have developed a qYN procedure that combines elements of signal detection theory and Bayesian adaptive inference to concurrently estimate a threshold associated with a d' level (rather than a percent yes level) and decision criterion (Lesmes et al., 2015).

In this study, we implemented the qVFM method with a YN light-detection task based on the qYN procedure. We will first describe the qVFM algorithm, and then computer simulations and a psychophysical experiment to validate the method in assessing the light-sensitivity VFM.

The qVFM method

The qVFM method consists of three major modules. In the global module, the shape of the VFM is modeled as a tilted elliptic paraboloid function (TEPF) with five parameters. The score of the VFM at each visual-field location represents a measure of visual function in that location. For example, in mapping light sensitivity, the score in each location of the VFM represents perceptual sensitivity (1/threshold) at a fixed d' level in that location. Together with a decision criterion (the sixth parameter) and a slope of the psychometric function (assumed to be fixed), the VFM can be used to predict light-detection probability in every single visual-field

location. Using Bayesian update and optimal stimulus selection (Kontsevich & Tyler, 1999; Lesmes et al., 2015), the qVFM updates the joint posterior distribution of the six parameters—that is, the shape of the VFM—based on subject's response in each trial.

The switch module evaluates the rate of information gain in the global module and determines when to switch to the local module. At the point of the switch, the module generates a prior distribution of measures of visual function in each visual-field location based on the posterior from the global module. For example, in mapping light sensitivity, the prior is over both perceptual sensitivity and decision criterion at each visual-field location.

Using the prior generated by the switch module, the local module provides assessment of visual function in each visual-field location via another Bayesian adaptive procedure that determines the order and stimulus of the test based on the relative information gain across locations. For example, in mapping light sensitivity, we used qYN (Lesmes et al., 2015) as the adaptive procedure to assess visual function at each visual-field location, and the expected information gain across all visual-field locations and stimulus intensity levels to determine the optimal stimulus in each trial.

Next, we describe all three modules in more detail.

The global module

In the global module, we first define a six-dimensional prior probability distribution, in which five of the dimensions correspond to the five parameters of the TEPF and one dimension represents the decision criterion in a YN task in light detection. The prior distribution, together with the slope of the psychometric function, completely specifies the probability of YN responses across all the light intensity levels and visual-field locations of all possible observers. We then define a two-dimensional stimulus space representing both the spatial locations and intensities of the stimuli. The optimal stimulus (location and light intensity) for the first test is determined based on the prior distribution. Bayes's rule is used to update the posterior distribution of the six parameters based on the observer's response. A new trial starts using the posterior from the previous trial as the prior. The procedure repeats until the switch module decides to switch to the local module.

Modeling the likelihood function of the VFM with a TEPF

In the global module, we model the VFM as a TEPF of spatial location (as shown in Figure 1) with five parameters: the central gain (e.g., sensitivity at the
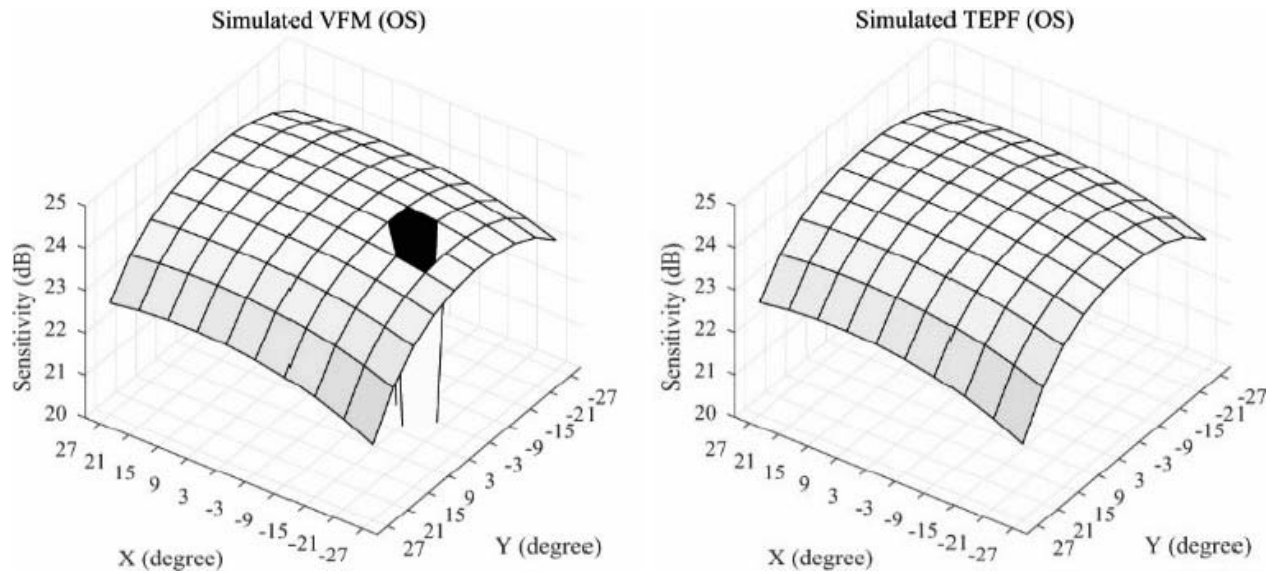
Figure 1. Simulated visual-field map and tilted elliptic paraboloid function. The left panel shows the visual-field map of a simulated left eye with a blind spot at (−15°, −3°). The right panel shows the corresponding tilted elliptic paraboloid function without the blind spot.

fovea), EPZ; (2) the bandwidth (latus rectum) in the horizontal direction, EPA, which describes the function's full width at half maximum (in octaves) in the horizontal direction of the visual field; the bandwidth in the vertical direction, EPB; the tilted level in the horizontal direction, SLA; and the tilted level in the vertical direction, SLB:

$$\tau(x, y) = \text{EPZ} - \left(\frac{x}{\text{EPA}}\right)^2 - \left(\frac{y}{\text{EPB}}\right)^2 + \text{SLA} \times x + \text{SLB} \times y. \quad (1)$$

In mapping light sensitivity, the height of the TEPF, $\tau(x,y)$, defines the light sensitivity (1/threshold) at a fixed $d' = 1.0$ level at visual-field location (x,y).

The d' psychometric function at each visual-field location (x,y)—that is, perceptual sensitivity for a given stimulus intensity s—is modeled as

$$d'(s, x, y) = \frac{\beta(s \times \tau(x, y))^\gamma}{\sqrt{(s \times \tau(x, y))^{2\gamma} + (\beta^2 - 1)}}, \quad (2)$$

where γ is the slope of the d' psychometric function and β is the asymptote of the function (Lesmes et al., 2015; Lu & Dosher, 2008). Plotted on log axes, this function is approximately linear over low to medium stimulus intensities and saturates at high intensities. Following the previous study (Lesmes et al., 2015), we have fixed γ (=2.1) and β (=5.0) in the current implementation of the qVFM.

In the YN detection task at a visual-field location (x,y), the probability of reporting target presence ("yes") is determined by both perceptual sensitivity and decision criterion. Based on signal detection theory (X. Gu & Green, 1994; Klein, 2001),

$$P(s, x, y) =$$
$$\int_{-\infty}^{+\infty} \phi(t - [d'(s, x, y) - \lambda(x, y)])\Phi(t)dt, \quad (3)$$

where $\phi()$ is the probability density function of a standard normal distribution function, $\Phi()$ is the cumulative probability density function of a standard normal distribution function, $d'(s,x,y)$ is the d' value associated with a stimulus with signal intensity s at visual-field location (x,y), and $\lambda(x,y)$ is the decision criterion at visual-field location (x,y). In the global module, a single λ is used across all visual-field locations. In the local module, $\lambda(x,y)$ is independent at each visual-field location.

In addition, we assume a fixed lapse rate ε for human observers (Klein, 2001; Lesmes et al., 2015; Wichmann & Hill, 2001):

$$P'(s, x, y) = \frac{1}{2}\varepsilon + (1 - \varepsilon)P(s, x, y), \quad (4)$$

where $P(s,x,y)$ is the psychometric function without lapse. In the qVFM method, ε was set to 0.03 (Lesmes et al., 2010; Wichmann & Hill, 2001).

Equation 4 is the likelihood function of the VFM that completely describes the probability of light detection across all visual-field locations and light intensity levels.

Setting the prior and stimulus space

Before running the qVFM procedure, a probability density function $p(\vec{\theta})$ is defined over the parameter space of the TEPF and decision criterion, where $\vec{\theta} = (EPZ, EPA, EPB, SLA, SLB, \lambda)$. Before any data collection (trial $t = 0$), the initial prior distribution $p_{t=0}(\vec{\theta})$ represents foreknowledge of model parameters. In addition, a stimulus space is defined that includes all possible stimulus locations (x,y) and stimulus intensities. With the prior and the stimulus space, we can compute the probability of detecting any stimulus in the visual field for all possible observers based on Equation 4.

Stimulus selection

In the qVFM, information is quantified by entropy, a measure of uncertainty associated with random variables. The method uses a one-step-ahead search strategy to determine the optimal stimulus in the next trial that would lead to the minimum expected entropy. It first computes the observer's response probability $P_{t+1}(r|s)$ in every possible stimulus condition in the next trial based on the current prior; the expected posterior probability distributions for all possible stimuli; and the expected entropy for each possible stimulus.

The entropy of the posterior is defined as

$$H_{t+1}(s,r) = -\sum_{\vec{\theta}} P_{t+1}\left(\vec{\theta}|s,r\right) \times \log\left(P_{t+1}(\vec{\theta}|s,r)\right), \quad (5)$$

where r represents the observer's response (yes or no) to a test with signal intensity s. The expected entropy after a trial with the signal intensity of stimulus s is calculated as a weighted sum of posterior entropy:

$$E[H_{t+1}(s,r)] = \sum_{r} H_{t+1}(s,r) \times P_{t+1}(r|s). \quad (6)$$

The stimulus with the minimum expected entropy is chosen for the next trial:

$$s_{t+1} = \arg\min_{s} E[H_{t+1}(s)]. \quad (7)$$

This is equivalent to maximizing the expected information gain, quantified as the entropy change between the prior and posterior (Kujala & Lukka, 2006; Lesmes et al., 2006).

Bayesian update

The prior distribution $p_t(\vec{\theta})$ in trial t is updated to the posterior distribution $p_t(\vec{\theta}|s,r_t)$ with the observer's

response r_t (yes or no) to a test with a stimulus s by Bayes's rule:

$$P_t\left(\vec{\theta}|s, r_t\right) = \frac{P\left(r_t|\vec{\theta}, s\right) P_t\left(\vec{\theta}\right)}{P_t(r_t|s)}, \quad (8)$$

where $\vec{\theta}$ represents the parameters of the VFM model and $p_t(\vec{\theta})$ is the prior probability function of $\vec{\theta}$. The probability of a response r_t in a given stimulus condition s, $p_t(r_t|s)$, is estimated by weighting the empirical response probability by the prior:

$$P_t(r_t|s) = \sum_{\vec{\theta}} \left[P\left(r_t|\vec{\theta}, s\right) P_t\left(\vec{\theta}\right) \right], \quad (9)$$

where $p(r_t|s,\vec{\theta})$ is the likelihood of observing response r_t given $\vec{\theta}$ and stimulus s. The posterior $p_t(\vec{\theta}|s,r_t)$ following trial t serves as the prior $p_{t+1}(\vec{\theta})$ in the next trial:

$$P_{t+1}\left(\vec{\theta}\right) = P_t\left(\vec{\theta}|s, r_t\right). \quad (10)$$

The means of the marginal posterior distributions are used to estimate the parameters of the qVFM model after each trial.

The switch module

Since TEPF modeling cannot detail the blind spot or other regional deficits on the VFM (as shown in Figure 1), it is necessary to switch from the global module to the local module for further VFM assessments.

In the global module, the expected information gain—computed as the difference between the entropy of the prior and the expected posterior distribution for each potential stimulus—is computed before each trial. The stimulus that would lead to the maximum amount of information gain is used in the next trial. Before each trial, the expected information gain is also used by the switch module to determine the switch point to the local module. Instead of using the maximum expected information gain, the switch module computes the total expected information gain (TEI) from the top 10% potential stimuli. In the beginning, the TEI is high in the global module. With increasing numbers of trials, the TEI is expected to gradually decrease as the method learns more about the parameters. As the learning saturates over trials, the trend of the TEI begins to flatten and may even reverse—that is, the TEI in trial $t + 1$ may be higher than that in previous trials. In the current implementation, the switch module compares the TEI in trial $t + 1$ with the average TEI of three previous trials, $t - 2$, $t - 1$, and t, to determine if the qVFM should switch to the local module. The switch happens when the TEI in

trial $t + 1$ is higher than the average TEI in trials $t - 2$, $t - 1$, and t.

Upon the switch, the switch module generates a prior distribution of visual function in each visual-field location based on the posterior of the parameters in the global module. Specifically, the posterior distribution is sampled repeatedly to generate the prior distribution for each visual-field location.

The local module

The setup of the local module is very similar to that of the global module except the following:

- Independent parameters in each visual-field location: Instead of using six parameters to model visual function across all visual-field locations, independent parameters are used to model visual function in each visual-field location. In mapping light sensitivity, $\tau(x, y)$ in Equation 2 is no longer described by Equation 1 but rather is independent in each location. In addition, the decision criterion $\lambda(x, y)$ is independent at each visual-field location as well.
- Independent priors and posteriors: Each visual-field location has its independent parameters and therefore independent priors and posteriors. The initial priors in the local module are generated by the switch module.
- Computing the information gain: The information gain of each location is computed independently, while in the global module the information gains are computed simultaneously across all visual-field locations based on the TEPF and the psychometric function with the bias criterion. Regardless of the dependency of the computations of information gains, optimal stimulus selection is always based on the total expected entropy across all the visual-field locations in both local and global modules. In other words, to select the next test location and stimulus intensity, the expected entropy from all visual-field locations are considered.

Stopping rules

In the current implementation, the qVFM procedure terminates after a fixed number of trials. Alternatively, it can stop after it achieves a certain defined objective (e.g., after reaching a criterion level of precision for either the parameters in qVFM or the light sensitivity across all visual-field locations).

Simulations

Methods

To evaluate the performance of the qVFM method, we simulated an observer with healthy vision who viewed testing target (a light disc) monocularly and was cued to report the presence or absence of a light target in 100 visual-field locations (Figure 1). The parameters of the simulated observer were chosen to approximate those of the observers in our psychophysical validation, with EPA = 81.0 ($°/\sqrt{dB}$), EPB = 41.1 ($°/\sqrt{dB}$), EPZ = 24.3 (dB), SLA = 0.020 (dB/°), SLB = 0.032 (dB/°), and $\lambda = 1.20$. Here, dB is calculated as $-10 \times \log_{10}$(luminance, in asb/10,000). The blind spot of the simulated left eye was at $(-15°, -3°)$ of the VF.

In the qVFM, the parameter space includes 20 linearly spaced EPA values (from $69.0°/\sqrt{dB}$ to $93.0°/\sqrt{dB}$), 20 linearly spaced EPB values (from $33.6°/\sqrt{dB}$ to $51.6°/\sqrt{dB}$), 32 linearly spaced EPZ values (from 16.3 to 25.0 dB), 15 linearly spaced SLA values (from -0.2 to 0.2 dB/°), 15 linearly spaced SLB values (from -0.17 to 0.23 dB/°), and 20 linearly spaced λ values (from 0.4 to 2.1). The broad parameter space ensures robust assessment of a wide range of populations and avoids effects of extreme values—the tendency to bias toward the center of the parameter space when the observer's true parameter values are close to the boundary of the space.

For each of the six qVFM parameters, the prior was defined by a hyperbolic secant (sech) function (King-Smith & Rose, 1997). For each qVFM parameter θ_i ($i = 1, 2, 3, 4, 5, 6$), the mode of the marginal prior $p(\theta_i)$ was defined by the best guess for that parameter based on a pilot study, $\theta_{i,\text{guess}}$, and the width was defined by the confidence in that guess, $\theta_{i,\text{confidence}}$:

$$P(\theta_i) = \text{sech}\left(\theta_{i,\text{confidence}} \times \left(\theta_i - \theta_{i,\text{guess}}\right)\right), \quad (11)$$

where

$$\text{sech(z)} = \frac{2}{e^z + e^{-z}}. \quad (12)$$

The priors were log-symmetric around $\theta_{i,\text{guess}}$, whose values for the respective parameters were EPA = 71.4 ($°/\sqrt{dB}$), EPB = 46.1 ($°/\sqrt{dB}$), EPZ = 24.5 (dB), SLA = 0.019 (dB/°), SLB = 0.048 (dB/°), and $\lambda = 1.16$. For $\theta_{i,\text{confidence}}$ of the respective parameters, the values were set to 0.67 for EPA, 1.26 for EPB, 1.93 for EPZ, 3.13 for SLA, 3.03 for SLB, and 2.68 for λ. The joint prior was defined as the normalized product of the marginal priors, which resulted in a weakly informative prior in the current study.

The stimulus space includes a 10×10 evenly spaced grid of visual-field locations ($60° \times 60°$) and log-linearly

spaced luminance values (from 10.2 to 25.0 dB): 60 luminance values in the global module and 120 in the local module.

We compared the performance of the full qVFM procedure, which has all three modules, with a reduced qVFM procedure that has only the local module (the qYN procedure in this study), a staircase procedure based on the 4-2 algorithm, and a staircase procedure based on the 4-2-1 algorithm, in 1,000 repeated simulations of 1,200 trials each.

The prior of the reduced qVFM was generated from the prior of the global module of the full qVFM. Therefore, the two methods were equated before the first trial. For simplification, we will use "the qYN method" to refer to the reduced qVFM method from this point on.

To minimize the bias and variability of the staircase procedure and give it the best chance, we matched the initial stimulus intensity in each location with the true sensitivity of the simulated observer. The step size in the 4-2 staircase algorithm is 4 dB at beginning and decreases to 2 dB after the first response reversal. In the 4-2-1 algorithm, the step size keeps decreasing to 1 dB after the second response reversal. The test on each location stops at the second response reversal in the 4-2 algorithm and the third response reversal in the 4-2-1 algorithm. The sensitivities are estimated with two methods, both used in clinical perimetry (Anderson & Patella, 1999; Weijland, Fankhauser, Bebie, & Flammer, 2004): the last seen stimulus intensity (referred to as the h method) and the average of the last seen and unseen stimulus intensities (referred to as the o method). The test procedure follows the "growth pattern" method, which begins at four primary points and grows to their neighboring points in a pseudorandom order.

Metrics of evaluation

Accuracy is a measure of how much the estimates deviate from the truth and precision is a measure of the variability of the estimates. We quantify accuracy using the root mean square error (RMSE) of the estimated sensitivities across all 100 visual-field locations. $\text{RMSE}^{\text{simulation}}$ after trial i can be calculated as

$$\text{RMSE}_i^{\text{simulation}} = \sqrt{\frac{\sum_k \sum_j \left(\tau_{ijk} - \tau_k^{\text{true}}\right)^2}{J \times K}}, \quad (13)$$

where τ_{ijk} is the estimated sensitivity at the kth visual-field location after trial i in the jth run, and τ_k^{true} is the true sensitivity at that location.

Two methods have been used to assess the precision of the qVFM procedure. The first is based on the standard deviation of repeated measures:

$$\text{SD}_i = \sqrt{\frac{\sum_k \sum_j \left(\tau_{ijk} - \text{mean}\left(\tau_{ijk}\right)\right)^2}{J \times K}}. \quad (14)$$

Another measure of precision is the average half width of the credible interval (HWCI) of the posterior distribution of the estimated sensitivities across visual-field locations. The 68.2% credible interval represents the range within which the actual value lies with 68.2% probability, whereas the confidence interval, the most popular index of precision, represents an interval that contains the true value of sensitivity for 68.2% of unlimited repetitions. Since researchers typically do not iterate an experiment many times for the same observer, the HWCI of the posterior distribution is a very important index of precision that can be obtained with a single run of the procedure (Hou et al., 2015).

To further compare the performance of all methods, we computed a number of global indices on the estimated VFMs from the simulations for each method: mean defect, loss variance, short-term fluctuation, and corrected loss variance (Flammer, Drance, Augustiny, & Funkhouser, 1985). The mean defect of the estimated sensitivities across all 100 visual-field locations after trial i is calculated as

$$\text{MD}_i = \frac{\sum_k \sum_j \left(\tau_k^{\text{true}} - \tau_{ijk}\right)}{J \times K}. \quad (15)$$

The loss variance is calculated as

$$\text{LV}_i = \frac{\sum_k \sum_j \left(\tau_{ijk} - \tau_k^{\text{true}} + \text{MD}_i\right)^2}{J \times (K-1)}. \quad (16)$$

The short-term fluctuation is calculated as

$$\text{SF}_i = \sqrt{\frac{\sum_k \sum_j \left(\tau_{ijk} - \text{mean}\left(\tau_{ijk}\right)\right)^2}{(J-1) \times K}}. \quad (17)$$

The corrected loss variance is calculated as

$$\text{CLV}_i = \text{LV}_i - \text{SF}_i^2. \quad (18)$$

Results

The estimated light-sensitivity VFMs of the simulated observer, obtained with the qVFM and qYN methods, are shown in Figure 2.

Figure 3 shows the corresponding RMSE, standard deviation, and average 68.2% HWCI of the estimated VFM in each visual-field location.

The average $\text{RMSE}^{\text{simulation}}$ of the estimated sensitivities from both the qVFM and qYN methods across all 100 visual-field locations started at 1.41 dB on the first trial. It decreased to 0.76 dB and 1.15 dB for the
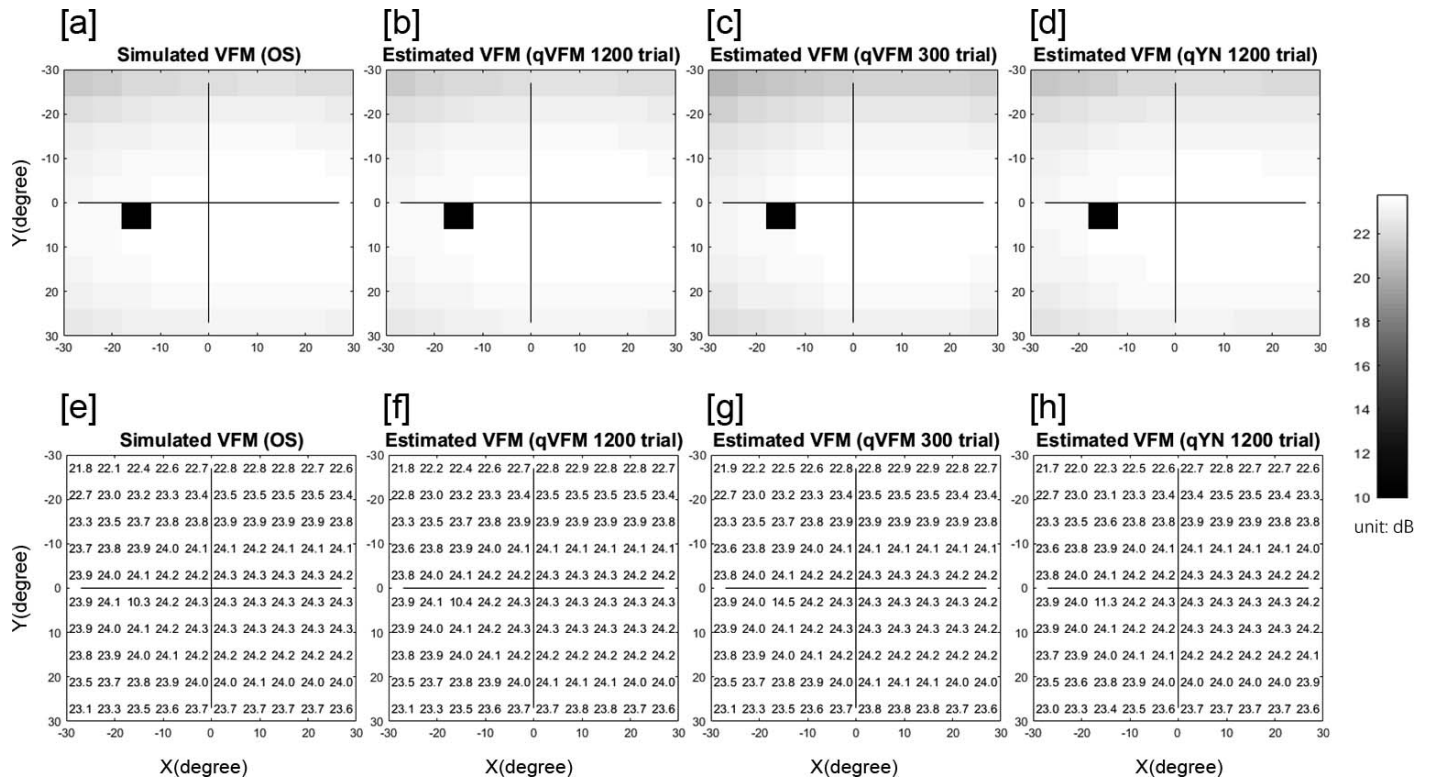
Figure 2. Simulation results I. (a, e) The true visual-field map (VFM) of a simulated observer (monocular). The estimated VFMs obtained with the qVFM method after (b, f) 1,200 trials and (c, g) 300 trials. (d, h) The estimated VFM obtained with the qYN method.

qVFM and qYN methods after the first 300 trials, and to 0.15 dB and 0.26 dB for the two methods after 1,200 trials (Figure 4a). The average RMSE^{simulation} of the staircase methods started at 0 dB and increased to 0.79 and 0.53 dB for the 4-2-1h and 4-2-1o algorithms and 1.39 and 0.74 dB for the 4-2h and 4-2o algorithms after 300 trials, and 0.92 and 0.65 dB for the 4-2-1h and 4-2-1o algorithms and 1.54 and 0.82 dB for the 4-2h and 4-2o algorithms after 1,200 trials (Figure 4a).

The standard deviations of the estimated sensitivities were 0.63 dB with the qVFM method and 0.75 dB with the qYN method after 300 trials, and decreased to 0.15 dB with the qVFM method and 0.24 dB with the qYN method after 1,200 trials (Figure 4b). The standard deviations of the staircase methods were 0.68 and 0.51 dB for the 4-2-1h and 4-2-1o algorithms, 0.91 and 0.68 dB for the 4-2h and 4-2o algorithms after 300 trials, and 0.62 dB for both the 4-2-1h and 4-2-1o algorithms and 0.73 dB for both the 4-2h and 4-2o algorithms after 1,200 trials (Figure 4b).

The 68.2% HWCI of the estimated sensitivities also decreased with trial number. It started at 2.12 dB for both the qVFM and qYN methods, and decreased to 0.20 dB with the qVFM method and 0.69 dB with the qYN method after the first 300 trials and to 0.15 dB with the qVFM method and

0.24 dB with the qYN method after 1,200 trials (in Figure 4c).

For the qVFM method, the switch from the global module to the local module occurred after between 30 and 63 trials, with the mean around 38 trials and a standard deviation of 6.8 trials. From Figure 4b and 4c, we can tell that the global module acted very efficiently in reducing random errors and uncertainties in the beginning of the qVFM method.

In characterizing spatial vision, the area under the log contrast sensitivity function is often used as a summary metric (Applegate et al., 2000; Applegate, Howland, Sharp, Cottingham, & Yee, 1998; Oshika, Klyce, Applegate, & Howland, 1999; Oshika, Okamoto, Samejima, Tokunaga, & Miyata, 2006; van Gaalen, Jansonius, Koopmans, Terwee, & Kooijman, 2009). Here in Figure 4d, we show the volume under the surface of the VFM (VUSVFM) to provide a summary metric of the entire visual field.

Figure 5a shows that the mean defects of both the qVFM and qYN methods did not drift much and converged to 0 dB with increasing trial number, while the Full Threshold method shows significant overall biased estimations across visual field. The loss variance and short-term fluctuation of the qVFM method both dropped quickly below those of the qYN method after several trials in the beginning, as the qYN method

Figure 3. Simulation results II. The (a–c) RMSE[simulation], (d–f) standard deviation, and (g–i) average 68.2% half width of the credible interval of the estimates (unit: dB) in each visual-field location for, respectively, the qVFM method after 1,200 trials, the qVFM method after 300 trials, and the qYN method after 1,200 trials.

exhibited vast fluctuations (Figure 5c and 5b). For the staircase methods, the loss variance all started at 0 dB^2, ending up at 0.39 dB^2 for both the 4-2-1h and 4-2-1o algorithms and 0.56 dB^2 for both the 4-2h and 4-2o algorithms (Figure 5c). The short-term fluctuations all increased in patterns similar to loss variance, ending up at 0.62 dB for both the 4-2-1h and 4-2-1o algorithms and 0.73 dB for both the 4-2h and 4-2o algorithms (Figure 5b). For both the qVFM and qYN methods, the corrected loss variance started at 1.95 dB^2 and dropped to less than 1 dB^2 after only 168 trials for the qVFM method, while the qYN method needed 241 trials to reach 1 dB^2.

In summary, the simulations showed that the staircase method exhibited overall positively biased estimates with high standard deviations. Although both the qVFM and qYN methods can reach high accuracy and precision in 1,200 trials, it took significantly fewer trials for the qVFM method. To achieve 1-dB accuracy (RMSE), qVFM needed only 213 trials, while qYN needed 352 trials. To achieve 1-dB precision (standard deviation and HWCI), qVFM
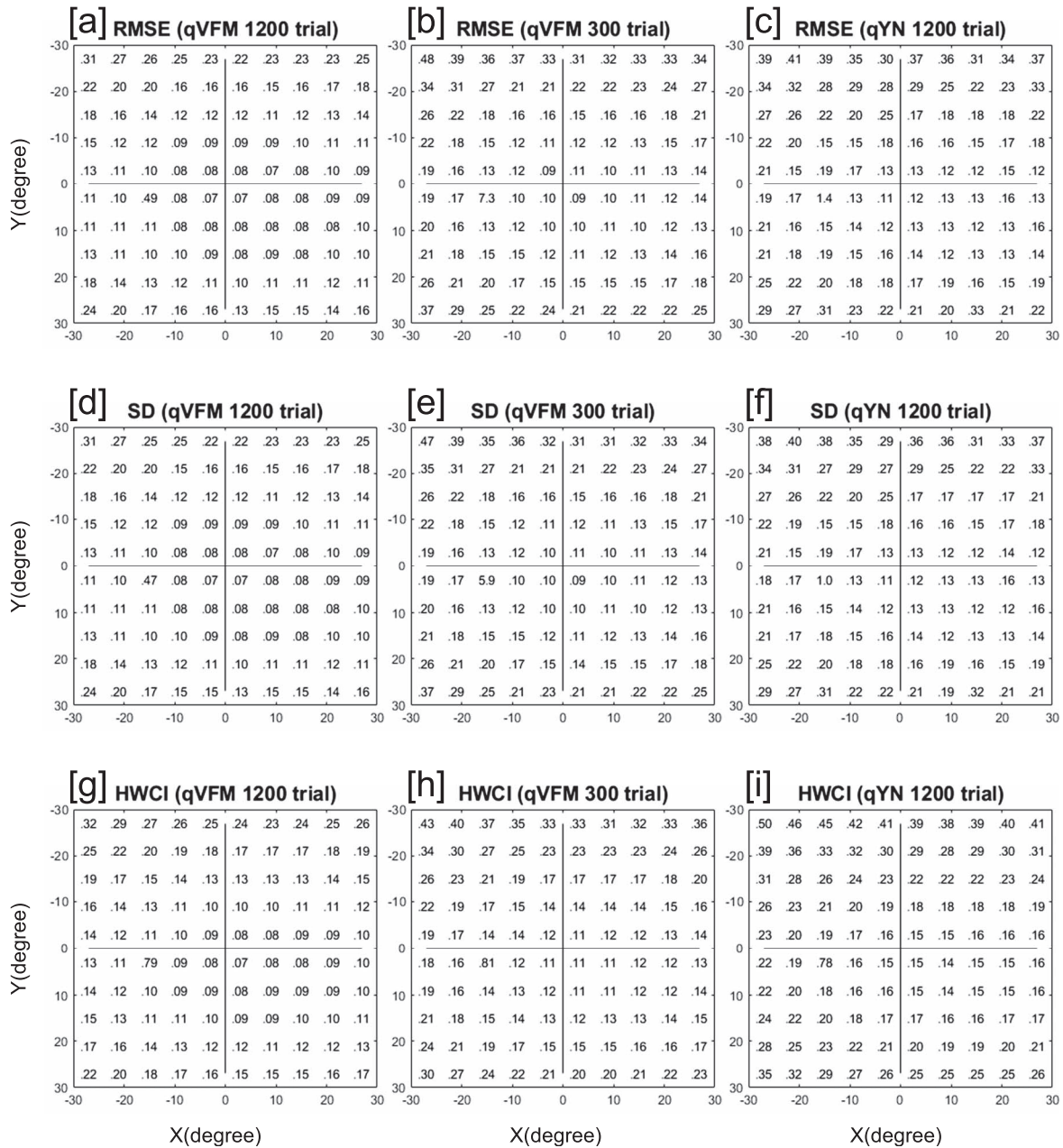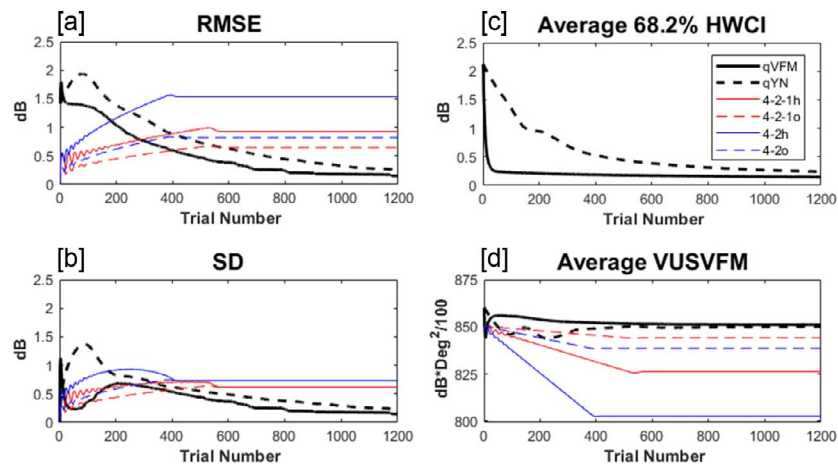
Figure 4. Simulation results III. (a) RMSE[simulation], (b) standard deviation, and (c) average 68.2% half width of the credible interval of the estimated sensitivities across 100 visual-field locations and 1,000 runs. (d) Average volume under the surface of the visual-field map across 1,000 runs. Results from the qVFM and qYN methods are shown in black solid and dashed lines. Results from the staircase 4-2-1 algorithm of h and o methods are shown in red solid and dashed lines. Results from the staircase 4-2 algorithm of h and o methods are shown in blue solid and dashed lines. The h method uses the last seen stimulus intensity as the estimated sensitivity, the o method uses the average of the last seen and unseen stimulus intensities as the estimated sensitivity.

used six and seven trials, while qYN used 155 and 152 trials.

Psychophysical validation

Methods

Apparatus

The psychophysical experiment was conducted on an IBM PC compatible computer running MATLAB (MathWorks, Natick, MA) programs with PsychToolbox extensions (Brainard, 1997; Pelli, 1997). The stimuli were displayed on a Samsung 55-in. monitor (UN55FH6030, Clear Motion Rate of 240 Hz) with a resolution of 1,920 × 1,080 pixels, a refresh rate of 60 Hz, and a background luminance of 31.5 asb. Subjects viewed the stimuli monocularly with natural pupil at a viewing distance of 30 cm in a dimly lighted room. A chin and forehead rest was used to minimize head movement during the experiment.
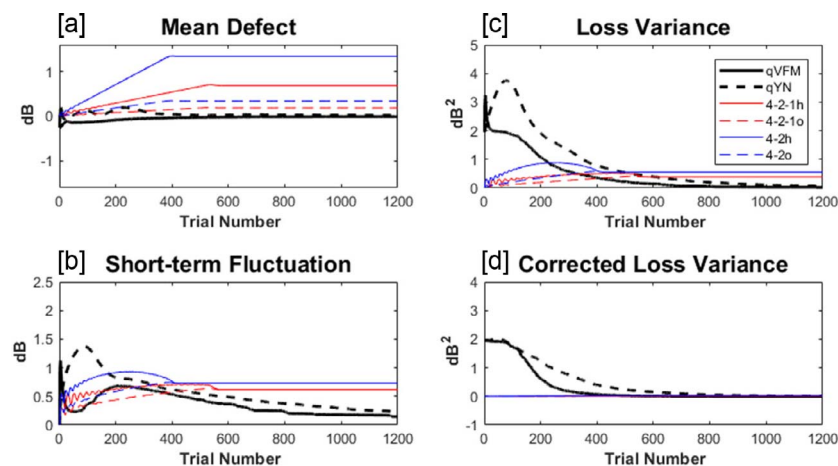
Figure 5. Simulation results IV. The global indices of the estimated visual-field maps: (a) mean defect, (b) short-term fluctuation, (c) loss variance, and (d) corrected loss variance. Results from the qVFM and qYN methods are shown in black solid and dashed lines. Results from the staircase 4-2-1 algorithm of h and o methods are shown in red solid and dashed lines. Results from the staircase 4-2 algorithm of h and o methods are shown in blue solid and dashed lines.
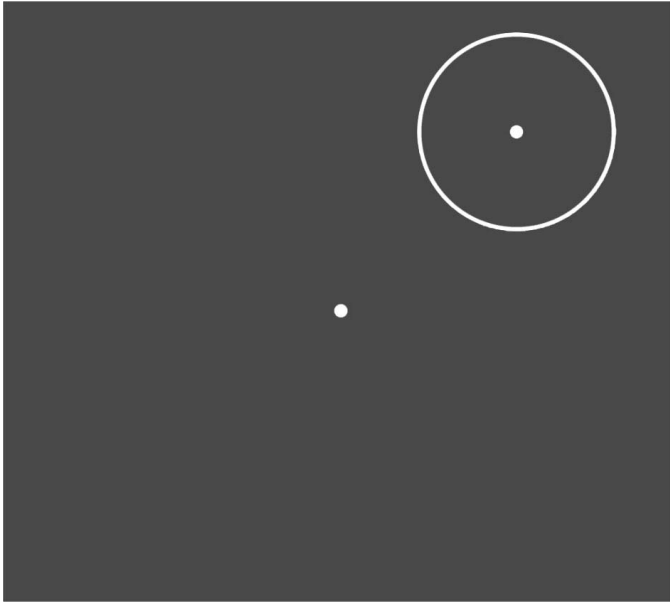
Figure 6. Illustration of the stimulus layout in the psychophysical experiment. Subjects were asked to fixate on the center of the display and report whether the target (a light disc with a diameter of 0.43° at one of the 100 visual-field locations) was present at the center of the location cue (circle) for 150 ms.

Stimuli and design

Consistent with the Goldmann size III stimulus (Dersu, Wiggins, Luther, Harper, & Chacko, 2006), the test target was a small light disc with a diameter of 0.43°. The luminance of the target varied between 31.5 and 950 asb, corresponding to sensitivity 25.0 to 10.2 dB, where $dB = -10 \times \log_{10}(\text{luminance, in asb}/10,000)$. Each trial contained a location cue—a circle with a diameter of 8.17° and a luminance of 77.4 asb. The target, if present, resided at the center of the location cue, and both were presented simultaneously for 150 ms in one of the 100 visual-field locations. All subjects confirmed afterwards that they were able to see the location cue in all testing trials. The intertrial interval was set to 1.2 s. A fixation dot was displayed in the center of the visual field throughout the whole session. An illustration of the stimulus layout is shown in Figure 6.

Procedure

Subjects were asked to report the target's presence or absence in the center of the location cue, with the luminance of the target adaptively adjusted in each trial. Each eye was tested in four sessions, each consisting of an independent 300-trial qVFM assessment and 300 serial qYN trials, with the two types of trials randomly mixed.

Participants

This study collected data from 12 eyes (six left, six right) of six subjects (four men, two women; 25–39 years old): the first author PX (Subject 1) and five others who were unaware of the research goals (Subjects 2–6).

Results

The estimated light-sensitivity VFMs of the 12 tested eyes from both the qVFM and qYN methods are shown in Figure 7 (Subject 1) and Figures A1 through A5 (Subjects 2–6).

The agreement between the estimated VFMs from qVFM and qYN was evaluated by the root mean square error (RMSE$^{\text{eyes}}$) of the estimated sensitivities across all 100 visual-field locations:

$$\text{RMSE}_i^{\text{eyes}} = \sqrt{\frac{\sum_l \sum_k \sum_j \left(\tau_{ijkl}^{\text{qVFM}} - \tau_{kl}^{\text{qYN}}\right)^2}{J \times K \times L}}, \quad (19)$$

where $\tau_{ijkl}^{\text{qVFM}}$ is the estimated sensitivity from the qVFM method in the kth visual-field location of eye l after i trials obtained in the jth session, and τ_{kl}^{qYN} is the estimated sensitivity from the qYN method in the kth visual-field location of eye l after 1,200 trials. The average RMSE$^{\text{eyes}}$ started at 2.47 dB on the first qVFM trial and decreased to 1.87 dB after 150 trials and to 1.38 dB after 300 trials across all test sessions and eyes (Figure 8a). The decreasing RMSE$^{\text{eyes}}$ with trial number suggests that the accuracy of qVFM increased with the number of trials.

The average 68.2% HWCI of the estimated sensitivities across all 12 eyes and 100 visual-field locations decreased from 2.28 dB before the first qVFM trial to 0.27 dB after 150 trials and 0.25 dB after 300 trials. The average 68.2% HWCI of the estimated sensitivities decreased from 2.28 dB before the first qYN trial to 1.05 dB after 150 trials, 0.69 dB after 300 trials, and 0.29 dB after 1,200 trials (as show in Figure 8b). These results suggest that the precision of the estimated sensitivities from the qVFM and qYN methods increased with trial number, and reached 1 dB in about 9 and 160 trials, respectively.

For the qVFM method, the switch from the global module to the local module occurred after between 26 and 54 trials, with the mean around 37 trials and a standard deviation of 5.7 trials across all 12 eyes, consistent with the simulations. The rapid convergence of the VFM estimation by the global module (the average 68.2% HWCI) is evident in Figure 8b.

Figure 8c presents the average estimated VUSVFM of the 12 eyes as a function of trial number for qVFM and qYN. The difference of the estimated VUSVFM
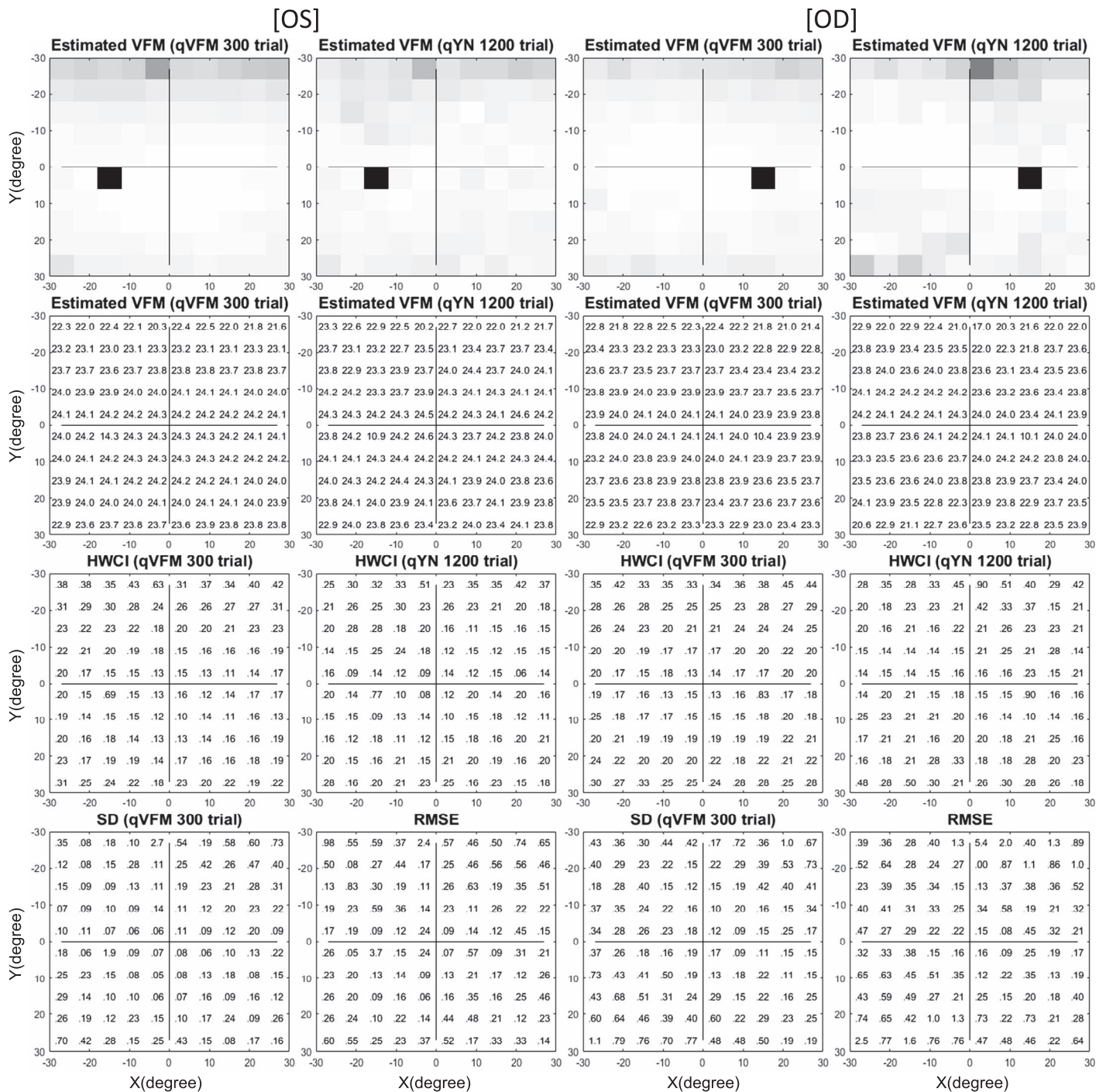
Figure 7. Experimental results I. Light-sensitivity visual-field maps (VFMs) in the left and right eye of Subject 1. The estimated VFMs are presented in the first row with achromatic color maps and in the second row with numerical values (unit: dB). For each visual-field location of the estimated VFM, the 68.2% half width of the credible interval is presented in the third row, and the standard deviations from four sessions of the qVFM method and the RMSE^{eyes} between the qVFM and qYN methods are presented in the fourth row. The results obtained from the qVFM and qYN methods are displayed in the odd and even columns, respectively.

between the two methods was less than 0.9% at 300 trials.

Test–retest reliability of the qVFM method is assessed through analysis of the four qVFM runs completed in four sessions. Figure 9a plots estimated

sensitivities (100 locations) of the paired qVFM runs from four independent sessions (2 random pairs of qVFM × 12 eyes × 100 locations = 2,400 data points). The average test–retest correlation for the paired VFM
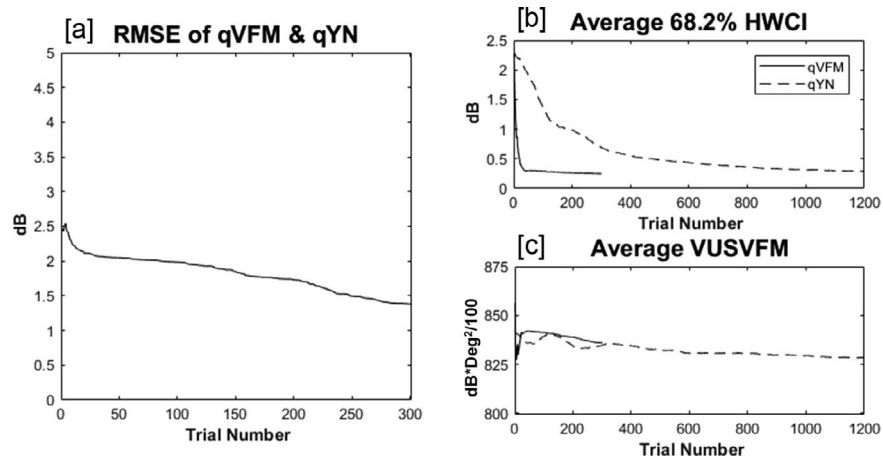
Figure 8. Experimental results II. (a) RMSE[eyes] of the estimated sensitivities from the qVFM method as a function of trial number, using estimated sensitivities from 1,200 qYN trials as the "truth." (b) Average 68.2% half width of the credible interval of the estimated sensitivities across 100 locations and 12 eyes. (c) Average volume under the surface of the visual-field map across 12 eyes. Results from the qVFM method are shown in solid lines, and results from the qYN method are shown in dashed lines.

estimate was 99.5% ± 0.06% (*SD*) across 100 runs of randomly sampled pairs.

Although test–retest correlations are widely reported as measures of test–retest reliability, they are not the most useful way to characterize method reliability or agreement (Bland & Altman, 1986). Figure 9b presents a Bland–Altman plot of the difference of the qVFM estimates between randomly sampled pairs of sessions against their respective means. The mean and standard deviation of the test–retest differences are −0.02 and 1.28 dB in Figure 9b. These results suggest that estimated sensitivities did not change much over the course of test sessions and that the test–retest difference between sessions agreed with the estimated RMSE[eyes]

(1.28 vs. 1.38 dB). The repeated runs of the qVFM procedure generated quite consistent results.

To demonstrate the convergence of the estimated VUSVFMs obtained with the qVFM method, Figure 9c presents the coefficient of variation of VUSVFM estimates as a function of trial number for each eye. The coefficient of variation, also known as relative standard deviation, is defined as the ratio of the standard deviation to the mean:

$$cv_i = \frac{\sigma_i}{\mu_i}, \quad (20)$$

where σ_i is the standard deviation of estimated VUSVFMs after trial i across four runs, and μ_i is the mean of the estimated VUSVFMs after trial i across
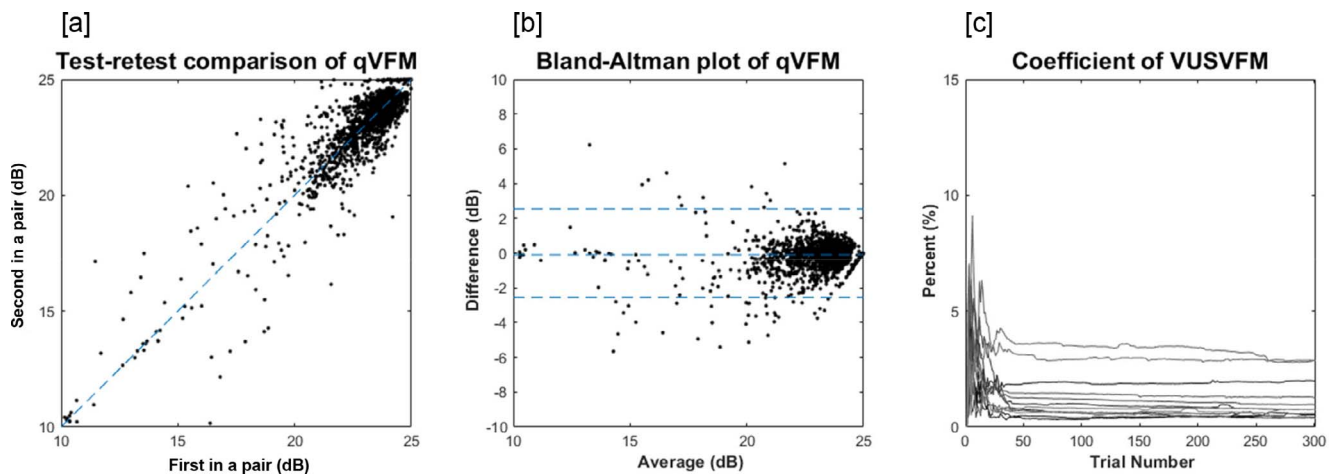
Figure 9. Experimental results III. (a) Test–retest comparison of the estimated sensitivities from repeated qVFM runs. (b) Bland–Altman plot for repeated qVFM runs. (c) Coefficient of variability of estimated volumes under the surface of the visual-field map (four runs each) as functions of trial number for all 12 tested eyes.

four runs. A consistent pattern, exhibited in each tested eye, is a decrease in variability with trial number: from approximately 8% after 15 trials to less than 3% after 300 trials.

Discussion

In this study, we developed a novel Bayesian adaptive testing framework, the qVFM method, that combines a global module for preliminary assessment of the VFM's shape and a local module for assessing individual visual-field locations. We implemented and validated the method in measuring light-sensitivity VFM. Simulations showed that the $RMSE^{simulation}$ and standard deviation of the estimated sensitivities after 1,200 trials were both 0.15 dB for the qVFM, and 0.26 and 0.24 dB for the qYN. To achieve 1-dB accuracy and 1-dB precision, on average, it took 213 and six or seven qVFM trials, respectively, and 352 and 152–155 qYN trials. Estimates of within-run variability (68.2% HWCIs) were comparable to cross-run variability (*SD*). For the subjects in the psychophysical experiment, the average HWCI of the qVFM estimates decreased from 2.28 dB on the first trial to 0.27 dB after 150 trials, and to 0.25 dB after 300 trials. The $RMSE^{eyes}$ of light-sensitivity estimates from the qVFM and qYN methods started at 2.47 dB on the first trial and decreased to 1.87 dB after 150 qVFM trials and to 1.38 dB after 300 trials. The qVFM provides an accurate, precise, and efficient mapping of light sensitivity. The method can be extended to map other visual functions, with potential clinical signals for monitoring vision loss, evaluating therapeutic interventions, and developing effective rehabilitation for low vision.

SAP, widely used in clinical vision, provides only a light-sensitivity map (Johnson et al., 2011). Other visual functions (e.g., visual acuity and contrast sensitivity function [CSF]) are typically measured only in one single spatial location, either in central vision or at a single peripheral location (Dorr et al., 2015; Hou et al., 2010; Qiu, Xu, Zhou, & Lu, 2007; Westheimer, 1965; Xu, Lu, Qiu, & Zhou, 2006; Xu et al., 2010). Vision tests in a single visual-field location cannot fully represent the typical usage of residual vision in ophthalmic patients' everyday activities (Bengtsson et al., 2005; Hodapp et al., 1993; Markowitz, 2006). To obtain a more comprehensive assessment of residual vision in patients, we must obtain other visual-function maps, including but not limited to visual acuity (Thompson et al., 1982; VA, 1965), CSF (Daitch & Green, 1969; Swanson, Malinovsky, et al., 2014), color vision (Carlow et al., 1976; Hart et al., 1984; Sample & Weinreb, 1990, 1992), motion sensitivity (Sperling & Lu, 1998), reading speed

(Ramulu et al., 2009; Yu et al., 2010), and crowding (Balas et al., 2009; Levi & Carney, 2009). It is extremely challenging or even impossible to obtain even a single one of these maps with conventional methods because of the tremendous amount of data collection required. The development of our hybrid adaptive procedures makes it possible to measure visual functions at multiple locations across the visual field.

Typically, visual functions are derived from individual thresholds in many conditions estimated with an adaptive procedure. Traditional adaptive procedures have been developed to estimate threshold and in a few cases the slope of the psychometric function for only a single stimulus condition at a time (Alcala-Quintana & Garcia-Perez, 2007; García-Pérez & Alcalá-Quintana, 2007; King-Smith & Rose, 1997). Although measurement of individual thresholds is efficient, these conventional methods do not take advantage of the relationship across conditions, and the demand of data collection is multiplied by the number of conditions.

Novel combination of global and local modules

The combination of the global and local modules to measure visual-function maps is a novel development in Bayesian adaptive testing techniques. Previous studies developed the qCSF method to accurately estimate the CSF in less than 100 trials (Dorr et al., 2015; Hou et al., 2010; Hou et al., 2015; Lesmes et al., 2010). In the qCSF method, the truncated log parabola with four parameters is used as the functional form of the CSF. The algorithm applies a one-step-ahead search, and an information-gain strategy, to efficiently estimate the four parameters of this functional form. With a 10-alternative forced-choice task (Hou et al., 2015), as few as 25 to 50 qCSF trials, distributed over 12 possible spatial frequencies, are sufficient for a general assessment of the CSF.

In this study, we adopted the method used in qCSF to develop the global module for preliminary assessment of the VFM's shape, which was modeled by a mathematical function, the TEPF, with five parameters. The adaptive Bayesian procedure was used to update estimates of the parameters of the model based on trial-by-trial performance. To provide detailed assessment in the individual visual-field locations, we introduced the local module using qYN (Lesmes et al., 2015) as the adaptive procedure, to assess visual function at each location with the expected information gain across all locations and stimulus intensity levels to determine the optimal stimulus in each trial. To determine the appropriate switch point from the global module to the local module, we evaluate the rate of information gain in the global module and switch to the

local module when the rate of information gain is lower than a criterion value. Further, we generate a prior distribution for the local module by sampling the posterior from the global module. Different global, local, and switch modules can be adopted and modified according to the specific applications of this framework.

This hybrid Bayesian adaptive test framework shows the capability of significantly reducing testing time for estimating visual function maps of normally-sighted individuals, laying the groundwork for future research on a wide range of eye diseases, including age-related macular degeneration (AMD), glaucoma, diabetic retinopathy, and retinitis pigmentosa.

Applying the qVFM method to estimate VFM in eye disease

We have evaluated the performance of the qVFM method on simulated observers with scotoma, cataract, glaucoma, and AMD. The results show that the qVFM method can provide an accurate, precise, and efficient assessment of vision loss in light-sensitivity maps.

In these simulations, the simulated glaucoma observer had affected peripheral vision, in which light sensitivity dropped to 2.5 times lower than that of the simulated normal observer outside of the central 6°. The simulated scotoma observer had three scotomas located at (−9°, 9°), (−9°, −9°), and (−9°, −15°) from the fovea. The simulated cataract observer had light sensitivity that was 2.2 dB lower compared to the simulated normal observer at each retinal location across the entire visual field. The simulated AMD observer had reduced light sensitivity across fovea vision in the central 6° of the visual field.

We sampled 100 visual-field locations (60° × 60°) and compared the performance of the qVFM with the qYN procedure that evaluated each location independently. The task for the simulated observers was to report a the presence or absence of a light disc, with its luminance adaptively adjusted on each trial. Simulated runs of 1,200 trials (for both qVFM and qYN) were used to compare the accuracy and precision of the methods.

For the simulated glaucoma observer, the RMSE of the qVFM and qYN estimates, started at 6.40 dB and became 1.52 dB and 3.13 dB after 300 trials and 0.65 dB and 0.73 dB after 1,200 trials, respectively. The average standard deviations of the qVFM and qYN estimates were 0.76 dB and 2.45 dB after 300 trials and 0.58 dB and 0.71 dB after 1,200 trials, respectively. For the simulated scotoma observer, the RMSE of the qVFM and qYN estimates started at 2.54 dB and became 1.42 dB and 1.71 dB after 300 trials and 0.29 dB and 0.37 dB after 1,200 trials, respectively. The average standard deviations of the qVFM and qYN estimates were 1.15

dB and 0.89 dB after 300 trials and 0.29 dB and 0.33 dB after 1,200 trials, respectively. For the simulated cataract observer, the RMSE of the qVFM and qYN estimates started at 4.63 dB and became 0.85 dB and 2.15 dB after 300 trials and 0.37 dB and 0.61 dB after 1,200 trials, respectively. The average standard deviations of the qVFM and qYN estimates were 0.65 dB and 1.77 dB after 300 trials and 0.36 dB and 0.60 dB after 1,200 trials, respectively. For the simulated AMD observer, the RMSE of the qVFM and qYN estimates started at 3.26 dB and became 1.98 dB and 2.23 dB after 300 trials and 0.30 dB and 0.52 dB after 1,200 trials, respectively. The average standard deviations of the qVFM and qYN estimates were 1.55 dB and 0.95 dB after 300 trials and 0.30 dB and 0.47 dB after 1,200 trials, respectively.

Comparison with staircase-based algorithms

Most of the existing algorithms for static automated perimetry are based on the staircase strategy (Weijland, Fankhauser, Bebie, & Flammer, 2004). In these algorithms, stimulus intensities are varied according to an up-and-down bracketing procedure in each location. The threshold values are estimated directly or scaled from the last seen stimulus intensity or the average of the last seen and unseen stimulus intensities in each location. In addition, the test procedures usually start from measuring thresholds at four primary points, one in each quadrant of the visual field, followed by measurement of thresholds in the rest of the visual field with initial values derived from the primary points.

Staircase-based perimetries are generally classified into conventional and reduced time-saving methods.

The conventional method with the Humphrey Field Analyzer is Full Threshold, which is currently regarded as the standard technique in static automated perimetry. With initial stimulus intensity levels determined from a normative data set, the stimulus intensity at each test location is varied in steps of 4 dB until the first response reversal occurs, and then subsequently varied in steps of 2 dB. The stimulus intensity of the last-seen presentation is taken as the final threshold estimate, after a second response reversal has occurred at a given location (Artes et al., 2002). The other conventional method implemented in Octopus perimeters uses a 4-2-1-dB staircase procedure, which further reduces the step size to 1 dB after two reversals. The mean value of the dimmest stimulus seen and the brightest stimulus not seen is defined as the threshold (Morales, Weitzman, & González de la Rosa, 2000).

The reduced method with the Humphrey Field Analyzer is FASTPAC. A step size of 3 dB is used, and the bracketing procedure stops after a single reversal. In half of the test locations, the first stimulus is 1 dB

brighter than the expected threshold, whereas for the other half the first stimulus is presented 2 dB dimmer than the expected threshold (Flanagan et al., 1993; O'Brien, Poinoosawmy, Wu, & Hitchings, 1994). Some reduced methods used in Octopus perimeters include the dynamic strategy and TOP (tendency-oriented perimetry; Johnson, 2002; Morales et al., 2000). In the dynamic strategy, the step size increases with the depth of visual defect and varies between 2 and 10 dB; the threshold is crossed only once. In TOP, each test location is assessed only once, and the subject's responses in neighboring locations are used to determine the threshold in each location (King, Taguri, Wadood, & Azuara-Blanco, 2002; McKendrick, 2005). Another algorithm, German Adaptive Threshold Estimation, compares the threshold obtained from the primary locations to the age-corrected normal hill of vision, then uses the smallest absolute deviation from the normal hill of vision to translate the values of the entire hill of vision. For the rest of the visual-field locations, testing starts slightly above the expected normal threshold and a 4-2-dB staircase is used to estimate the threshold after two reversals (Schiefer et al., 2009).

In our study, staircase procedures with both the 4-2 and 4-2-1 algorithms were implemented, with the initial stimulus intensity at each location matched with the true threshold of the simulated observer. The results show that staircase algorithm used in conventional procedures can have large biases and variabilities.

Comparison with SITA algorithms

The SITA algorithm, developed by Bengtsson, Olsson, Heijl, and Rootzén (1997), reduces the test time through more efficient threshold estimation based on Bayesian principles, whereas the stimulus intensities are varied according to simple staircase rules (similar to those of the Full Threshold strategy). It takes advantage of age-corrected healthy and glaucomatous prior probability distributions of threshold values to calculate the Bayesian posterior probability distributions based on participant's responses. The maximum posterior estimate is used to determine the best classification (healthy or glaucomatous) and calculate the threshold value in each visual-field location. The threshold value calculated in real time is used to determine when testing stops at each location.

The SITA family, including SITA Standard, SITA Fast, and SITA Faster, all share a similar framework, but with slight differences in the following features:

- Stimulus intensities: Whereas SITA Standard and SITA Fast both start the test sequence with a 25-dB stimulus intensity at the primary test locations, SITA Faster uses age-corrected healthy threshold

levels as initial stimulus intensity at the primary test locations, reducing the number of stimulus presentations for most eyes.
- Step size: Whereas SITA Standard uses 4-2-dB staircases, SITA Fast and SITA Faster alter stimulus intensities in 4-dB steps with one reversal in all test locations, except in the four primary locations in SITA Fast (Heijl et al., 2019).
- Priors: The prior threshold distribution in each test location is important for test efficiency. Whereas the priors in SITA Standard and SITA Fast are based on threshold distributions of healthy subjects obtained with the original Full Threshold method, SITA Faster uses threshold distributions of healthy subjects obtained with SITA Fast (Bengtsson, Heijl, & Olsson, 1998; Heijl et al., 2019).
- Stop rules: The test sequence of SITA Standard stops when the measurement error of the test locations reaches the error-related factor cutoff. In SITA Fast and SITA Faster, testing stops earlier with an increased cutoff value and therefore lower test accuracy (Bengtsson, Olsson, Heijl, & Rootzén, 1997; Heijl et al., 2019).
- Blind spot: Blind-spot catch trials are replaced by use of the Humphrey gaze tracker in SITA Faster. Retesting at perimetrically blind points, false-negative catch trials, and stimulus timing are slightly different in the SITA family (Bengtsson, Heijl, & Olsson, 1998; Bengtsson, Olsson, Heijl, & Rootzén, 1997; Heijl et al., 2019).

Although it is also based on a Bayesian adaptive testing framework, the qVFM method is quite different from the SITA algorithm.

Stimulus selection in the SITA family follows the conventional up-down staircase algorithm in each location with a step size equal or larger than 2 dB, and the test procedure follows the "growth pattern" procedure across the visual field (Bengtsson, Olsson, Heijl, & Rootzén, 1997). In the qVFM method, the step size in the stimulus space grid can be as small as 0.12 dB. Both global and local modules use the one-step-ahead search strategy across the entire visual field to determine the optimal stimulus in the next trial that would lead to the minimum expected entropy, equivalent to maximizing information gain on the next trial. More precise stimulus intensity and location selection in qVFM potentially lead to more accurate threshold estimation.

The frequency-of-seeing curve (a YN psychometric function) in SITA is not adjusted with the observer's decision criterion (Bengtsson, Olsson, Heijl, & Rootzén, 1997). Previous studies have shown that conventional YN threshold estimates exhibit approximately 25%–50% more variability (i.e., standard deviation) than criterion-free forced-choice threshold estimates (King-Smith, Grigsby, Vingrys, Benes, & Supowit, 1994; McKee, Klein, & Teller, 1985). In this

study, we implemented the qVFM method based on the qYN procedure, which combines elements of signal detection theory and Bayesian adaptive inference to concurrently estimate thresholds associated with a d' level (rather than a percent yes level) and decision criterion (Lesmes et al., 2015). Our study shows that this method can deliver criterion-free thresholds on a light-detection YN perimetric task with significant improvement in performance.

The prior distribution for SITA is optimized only for healthy and glaucoma observers, not for people with other eye diseases (Bengtsson & Heijl, 1998). In qVFM, the global module is an individualized test that takes into account the population properties in its prior but continues to optimize the test for each individual, including those with abnormal visual-field maps. Because the method is completely systematic and not oriented toward any particular pathological patterns, qVFM is not limited to a specific eye disease.

Because locations with defective visual sensitivity tend to appear in clusters in glaucomatous visual fields, the prior threshold distribution in each test location in SITA is calculated with interlocation correlations (Bengtsson, Olsson, Heijl, & Rootzén, 1997). Such correlations may not be present in other eye diseases. In qVFM, the local module is used to estimate visual threshold in each visual-field location independently, making it possible to detect steep visual-sensitivity changes in the visual field, such as scotoma resulting from optic neuropathies or visual-field islands in retinitis pigmentosa.

Since the age-corrected healthy and glaucomatous priors need enormous data collection to generate, SITA's development has so far been focused on light-sensitivity maps with Goldmann size III stimulus for the 30-2, 24-2, and 10-2 test patterns in the Humphrey Field Analyzer (Bengtsson, Olsson, Heijl, & Rootzén, 1997; Phu, Khuu, Zangerl, & Kalloniatis, 2017). Whereas the test area in SITA is limited to the central 30° of the visual field, with less than 76 test locations, qVFM can map larger areas of the visual field with different types of stimulus and a flexible number of test locations, without restrictions from prior knowledge as are found in SITA.

Ways to inform priors

A proper informative prior can further speed up the estimation process (Baek et al., 2016; H. Gu et al., 2016; Hou et al., 2010; Kim, Pitt, Lu, Steyvers, & Myung, 2014; Lesmes et al., 2006; Lesmes et al., 2010). The prior in qVFM can be informed by prior knowledge obtained in four different ways:

- Structure-based prior: We can use the structural images such as fundus images or OCT SLO

(OSLO) images (Landa, Rosen, Garcia, & Seiple, 2010; Menke, Sato, Van De Velde, & Feke, 2006; Okada et al., 2006) to localize scotomas, anatomic fovea, and preferred retinal locus and inform the prior in qVFM.
- Prior derived from statistical or machine learning: Statistical and machine-learning algorithms can be used to classify patients (Kononenko, 2001) and derive informative priors for different types of patients.
- Prior derived from the hierarchical adaptive approach: A hierarchical Bayesian extension of qVFM can provide a judicious way to exploit two complementary schemes of inference (with past and future data) to achieve even greater accuracy and efficiency in information gain (H. Gu et al., 2016; Kim et al., 2014). In this approach, each incoming subject is assigned to several possible patient categories with probabilities. Each category of patients has its own learned prior. The hierarchical qVFM simultaneously updates the probabilities of patient classification and the VFM throughout the testing process, and updates the priors of the categories after testing each new subject.
- Priors informed by other VFMs: For each new VFM (e.g., VFM of visual acuity), one can generate a prior using results from a previous VFM (e.g., VFM of light sensitivity).

Alternative algorithms and methods

In the current implementation of qVFM, we evaluated the rate of information gain in the global module and decided to switch to the local module when the rate of information gain was lower than a criterion value. Alternatively, the switching point can be practiced with other methods, such as comparison of the information gain from the global and local modules, the convergence of parameters of visual function in the global module, or a fixed trial number for prediagnosed eye diseases. Constrained by the number of computations required to maintain and update both the global and local modules, we implemented only a one-time switch from the global module to the local module in this study. It is possible to combine information from both modules in each trial with additional computational resources. Instead of a switch module, one can construct a communication module that provides constraints from both the top-down (global module) and the bottom-up (local module) views of the visual field.

In the current implementation of the qVFM method, the optimal test stimulus is selected based on maximum expected information gain, equivalent to the minimum

expected entropy, in each trial (Baek et al., 2016; Kujala & Lukka, 2006). Other metrics on information gain can be used, such as Fisher information, mutual information, or measures of the variability of the posterior distribution. Stimulus selection can also be practiced with other methods beyond the one-step-ahead strategy, such as multiple-steps-ahead search based on dynamic programming (Kim, Pitt, Lu, & Myung, 2017) or "region growing" techniques used in Humphrey and Octopus perimeters (Dua, Acharya, & Ng, 2011; Lee, 2019).

In this study, an exhaustive search algorithm was used. More advanced sampling methods, such as Markov-chain Monte Carlo, can be used to more efficiently search the stimulus space (Kujala & Lukka, 2006).

The qVFM procedure terminates after a fixed number of trials in this study. Alternatively, it can stop after it achieves a certain defined objective (e.g., after reaching a criterion level of precision). The dynamic stopping rule has been reported to be more efficient than the fixed-length procedure, especially in experiments with short runs (Alcala-Quintana & Garcia-Perez, 2007; Tanner, 2008).

Slope of psychometric function

Chauhan, Tompkins, LeBlanc, and McCormick, (1993) found that the slope of the frequency-of-seeing curve (a YN psychometric function) was highly correlated with the threshold or threshold deviation in conventional perimetry (Chauhan et al., 1993). The correlation was even higher for healthy subjects than those with suspected or actual glaucoma.

Due to the well-known difficulty and large number of trials required in estimating the slope of the psychometric function, a fixed-slope assumption is often used in parametric adaptive test procedures (Lesmes et al., 2015; Lu & Dosher, 2013; Watson & Pelli, 1983). King-Smith et al. (1994) found that when the fixed slope in the adaptive procedure was twice the real slope, the mismatch caused only a modest reduction in performance: After 20 trials with the mismatched slope, the overall weighted error was comparable to that after 17 trials with the matched slope, and the average bias was relatively small compared to the overall error of the threshold estimate. The authors suggest that, if optimal conditions are required for the study of subjects with impaired vision, it would probably be advantageous to use a relatively low slope value.

Another recent study (Lu, Zhao, Lesmes, Dorr, & Bex, 2019) shows that even under mismatched slope conditions, Bayesian adaptive methods with a fixed slope can generate unbiased threshold estimates for certain d' performance levels. The results provide a theoretical basis to use psychometric functions with fixed slopes in parametric Bayesian adaptive procedures.

Uncertainty and decision criterion

In SAP, the signal location is not known in advance; the observer cannot isolate the relevant location, and must monitor all locations. Such uncertainty lowers performance levels and leads to underestimated perceptual sensitivities (Lu & Dosher, 2013; Palmer, Ames, & Lindsey, 1993; Palmer, Verghese, & Pavel, 2000; Shaw, 1980). The test results can only be interpreted with complicated models based on signal detection theory that consider the responses from many detectors, one at each spatial location (Lu & Dosher, 2013). In this study, we used a circle to cue the target location and eliminated its location uncertainty. This greatly simplified the signal detection model.

Integrating qVFM with other assessments

The qVFM procedure can be integrated with a fundus camera or OCT SLO (OSLO) system to provide both structural and functional assessment of the visual field (Landa et al., 2010; Menke et al., 2006; Okada et al., 2006). Eye tracking can be integrated into a qVFM system to monitor fixation stability and measure subject's responses to the stimuli (Murray et al., 2009; Murray et al., 2013). Deep-learning networks can be used to learn the spatiotemporal VFM changes and generate predictions for disease progression (Wen et al., 2019).

Developing methodology for measuring VFM for a number of visual functions can allow us to analyze the relationships among all the metrics, such as thresholds or sensitivities of visual acuity, CSF, color, stereovision, temporary frequency, motion sensitivity, reading speed, and crowding maps. It would be helpful to evaluate and model the relationships between these metrics and performance in everyday visual tasks, and to identify the core metrics of functional vision in patients.

Conclusion

In this study, we developed a solution to address the major technological challenges in assessing VFM. Existing VFM assessment methods are too time consuming and unstable for clinical applications. The qVFM method based on the Bayesian adaptive testing framework will allow us to thoroughly characterize residual vision of ophthalmic patients and identify the core metrics of functional vision. The general framework

developed in this study can be potentially translated into clinical practice to examine visual deficits in patients, and extended to examine a broad range of eye diseases such as age-related macular degeneration, glaucoma, diabetic retinopathy, and retinitis pigmentosa.

Keywords: Bayesian adaptive testing, automated perimetry, visual-field map, peripheral vision, light sensitivity

Acknowledgments

Commercial relationships: PX, LAL, DY, and ZLL own intellectual-property rights in the qVFM technology. LAL and ZLL have equity interest in Adaptive Sensory Technology, Inc. LAL holds employment at Adaptive Sensory Technology, Inc.
Corresponding author: Zhong-Lin Lu.
Email: zhonglin@nyu.edu.
Address: Center for Neural Science and Department of Psychology, New York University, New York, NY, USA.

References

Advanced Glaucoma Intervention Study Investigators. (1994). Advanced Glaucoma Intervention Study: 2. Visual field test scoring and reliability. *Ophthalmology*, *101*(8), 1445–1455, https://doi.org/10.1016/S0161-6420(94)31171-7.

Alcala-Quintana, R., & Garcia-Perez, M. A. (2007). A comparison of fixed-step-size and Bayesian staircases for sensory threshold estimation. *Spatial Vision*, *20*(3), 197–218.

Applegate, R. A., Hilmantel, G., Howland, H. C., Tu, E. Y., Starck, T., & Zayac, E. J. (2000). Corneal first surface optical aberrations and visual performance. *Journal of Refractive Surgery*, *16*(5), 507–514.

Applegate, R. A., Howland, H. C., Sharp, R. P., Cottingham, A. J., & Yee, R. W. (1998). Corneal aberrations and visual performance after radial keratotomy. *Journal of Refractive Surgery*, *14*(4), 397–407.

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 & Visual Science*, *43*(8), 2654–2659.

Ashimatey, B. S., & Swanson, W. H. (2016). Between-subject variability in healthy eyes as a primary source of structural–functional discordance in patients with glaucoma. *Investigative Ophthalmology & Visual Science*, *57*(2), 502–507.

Aulhorn, E., & Harms, H. (1972). Visual perimetry. In M. Alpern, E. Aulhorn, H. B. Barlow, E. Baumgardt, H. R. Blackwell, D. S. Blough,... L. M. Hurvich (Eds.), *Visual psychophysics* (pp. 102–145), Berlin & Heidelberg, Germany: Springer, https://doi.org/10.1007/978-3-642-88658-4_5.

Baek, J., Lesmes, L. A., & Lu, Z.-L. (2014). Bayesian adaptive estimation of the sensory memory decay function: The quick partial report method. *Journal of Vision*, *14*(10):157, https://doi.org/10.1167/14.10.157. [Abstract]

Baek, J., Lesmes, L. A., & Lu, Z.-L. (2016). qPR: An adaptive partial-report procedure based on Bayesian inference. *Journal of Vision*, *16*(10):25, 1–23, https://doi.org/10.1167/16.10.25. [PubMed] [Article]

Balas, B., Nakano, L., & Rosenholtz, R. (2009). A summary-statistic representation in peripheral vision explains visual crowding. *Journal of Vision*, *9*(12):13, 1–18, https://doi.org/10.1167/9.12.13. [PubMed] [Article]

Bengtsson, B., Olsson, J., Heijl, A., & Rootzén, H. (1997). A new generation of algorithms for computerized threshold perimetry, SITA. Acta Ophthalmologica Scandinavica, *75*(4), 368–375. https://doi.org/10.1111/j.1600-0420.1997.tb00392.x

Bengtsson, B., & Heijl, A. (1998). Evaluation of a new perimetric threshold strategy, SITA, in patients with manifest and suspect glaucoma. *Acta Ophthalmologica*, *76*(3), 268–272.

Bengtsson, B., Heijl, A., & Agardh, E. (2005). Visual fields correlate better than visual acuity to severity of diabetic retinopathy. *Diabetologia*, *48*(12), 2494–2500.

Birch, M. K., Wishart, P. K., & O'Donnell, N. P. (1995). Determining progressive visual field loss in serial Humphrey visual fields. *Ophthalmology*, *102*(8), 1227–1235.

Bland, J. M., & Altman, D. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *The Lancet*, *327*(8476), 307–310.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, *10*(4), 433–436.

Broadway, D. C. (2012). Visual field testing for

glaucoma—A practical guide. *Community Eye Health*, 25(79–80), 66–70.

Brusini, P., Salvetat, M. L., Parisi, L., & Zeppieri, M. (2005). Probing glaucoma visual damage by rarebit perimetry. *British Journal of Ophthalmology*, 89(2), 180–184.

Caprioli, J. (1991). Automated perimetry in glaucoma. *American Journal of Ophthalmology*, 111(2), 235–239.

Carlow, T. J., Flynn, J. T., & Shipley, T. (1976). Color perimetry. *Archives of Ophthalmology*, 94(9), 1492–1496.

Chauhan, B. C., House, P. H., McCormick, T. A., & LeBlanc, R. P. (1999). Comparison of conventional and high-pass resolution perimetry in a prospective study of patients with glaucoma and healthy controls. *Archives of Ophthalmology*, 117(1), 24–33.

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 & Visual Science*, 40(3), 648–656.

Chauhan, B. C., Tompkins, J. D., LeBlanc, R. P., & McCormick, T. A. (1993). Characteristics of frequency-of-seeing curves in normal subjects, patients with suspected glaucoma, and patients with glaucoma. *Investigative Ophthalmology & Visual Science*, 34(13), 3534–3540.

Daitch, J. M., & Green, D. G. (1969). Contrast sensitivity of the human peripheral retina. *Vision Research*, 9(8), 947–952.

Dersu, I., Wiggins, M. N., Luther, A., Harper, R., & Chacko, J. (2006). Understanding visual fields, part I: Goldmann perimetry. *Journal of Ophthalmic Medical Technology*, 2, 1–10.

Dixon, W. J., & Mood, A. M. (1948). A method for obtaining and analyzing sensitivity data. *Journal of the American Statistical Association*, 43(241), 109–126.

Dorr, M., Wille, M., Viulet, T., Sanchez, E., Bex, P. J., Lu, Z.-L., & Lesmes, L. A. (2015). Next-generation vision testing: The quick CSF. *Current Directions in Biomedical Engineering*, 1(1), 131–134.

Drance, S. M., & Anderson, D. R. (1985). *Automatic perimetry in glaucoma: A practical guide*. Orlando, FL: Grune & Stratton.

Anderson, D. R., & Patella, V. M. (1999). *Automated static perimetry*. St. Louis, MO: Mosby.

Dua, S., Acharya, R., & Ng, E. Y. K. (2011). *Computational analysis of the human eye with applications*. Hackensack, NJ: World Scientific.

Easterbrook, M. (1984). The use of Amsler grids in early chloroquin retinopathy. *Ophthalmology*, 91(11), 1368–1372.

Fink, W., & Sadun, A. A. (2004). Three-dimensional computer-automated threshold Amsler grid test. *Journal of Biomedical Optics*, 9(1), 149–154.

Flammer, J., Drance, S. M., Augustiny, L., & Funkhouser, A. (1985). Quantification of glaucomatous visual field defects with automated perimetry. *Investigative Ophthalmology & Visual Science*, 26(2), 176–181.

Flanagan, J. G., Wild, J. M., & Trope, G. E. (1993). Evaluation of FASTPAC, a new strategy for threshold estimation with the Humphrey Field Analyzer, in a glaucomatous population. *Ophthalmology*, 100(6), 949–954.

Frisén, L. (1992). High-pass resolution perimetry: Evidence for parvocellular neural channel dependence. *Neuro-Ophthalmology*, 12(4), 257–264.

Frisén, L. (1993). High-pass resolution perimetry. *Documenta Ophthalmologica*, 83(1), 1–25.

García-Pérez, M. A., & Alcalá-Quintana, R. (2007). Bayesian adaptive estimation of arbitrary points on a psychometric function. *British Journal of Mathematical and Statistical Psychology*, 60(1), 147–174.

Goldmann, H. (1945a). Ein selbstregistrierendes Projektionskugelperimeter. *Ophthalmologica*, 109(2–3), 71–79.

Goldmann, H. (1945b). Grundlagen exakter perimetrie. *Ophthalmologica*, 109(2–3), 57–70.

Gu, H., Kim, W., Hou, F., Lesmes, L. A., Pitt, M. A., Lu, Z.-L., & Myung, J. I. (2016). A hierarchical Bayesian approach to adaptive vision testing: A case study with the contrast sensitivity function. *Journal of Vision*, 16(6):15, 1–17, https://doi.org/10.1167/16.6.15. [PubMed] [Article]

Gu, X., & Green, D. M. (1994). Further studies of a maximum-likelihood yes–no procedure. *The Journal of the Acoustical Society of America*, 96(1), 93–101.

Haley, M. J. (1986). *The field analyzer primer*. San Leandro, CA: Allergan Humphrey.

Harms, H. (1952). Die praktische Bedeutung quantitativer Perimetrie. *Klin Monbl Augenheilkd Augenarztl Fortbild*, 121(6), 683–692.

Hart, W. M., Hartz, R. K., Hagen, R. W., & Clark, K. W. (1984). Color contrast perimetry. *Investigative Ophthalmology & Visual Science*, 25(4), 400–413.

Heijl, A. (1977). Computer test logics for automatic perimetry. *Acta Ophthalmologica*, 55(5), 837–853.

Heijl, A. (1985). The Humphrey Field Analyzer, construction and concepts. In A. Heijl & E. L. Greve (Eds.), *Sixth International Visual Field*

Symposium: Santa Margherita Ligure, May 27–31, 1984 (pp. 77–84), Dordrecht, the Netherlands: Springer Netherlands, https://doi.org/10.1007/978-94-009-5512-7_10.

Heijl, A., Lindgren, A., & Lindgren, G. (1988). Interpoint correlations of deviations of threshold values in normal and glaucomatous visual fields. *Perimetry Update*, *89*, 177–183.

Heijl, A., Lindgren, A., & Lindgren, G. (1989). Test-retest variability in glaucomatous visual fields. *American Journal of Ophthalmology*, *108*(2), 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*(11), 1544–1549.

Heijl, Anders., Patella, V. M., Chong, L. X., Iwase, A., Leung, C. K., Tuulonen, A., Bengtsson, B. (2019). A New SITA Perimetric Threshold Testing Algorithm: Construction and a Multicenter Clinical Study. *American Journal of Ophthalmology*, *198*, 154–165. https://doi.org/10.1016/j.ajo.2018.10.010

Higgitt, A., & Smith, R. (1955). Reading test in glaucoma. *The British Journal of Ophthalmology*, *39*(2), 103–108.

Hodapp, E., Parrish, R. K., & Anderson, D. R. (1993). *Clinical decisions in glaucoma*. St. Louis, MO: Mosby.

Hood, D. C., Odel, J. G., & Winn, B. J. (2003). The multifocal visual evoked potential. *Journal of Neuro-Ophthalmology*, *23*(4), 279–289.

Hou, F., Huang, C., Lesmes, L. A., Feng, L., Tao, L., Zhou, Y., & Lu, Z.-L. (2010). qCSF in clinical application: Efficient characterization and classification of contrast sensitivity functions in amblyopia. *Investigative Ophthalmology & Visual Science*, *51*(10), 5365–5377.

Hou, F., Lesmes, L. A., Bex, P. J., Dorr, M., & Lu, Z.-L. (2015). Using 10AFC to further improve the efficiency of the quick CSF method. *Journal of Vision*, *15*(9):2, 1–18, https://doi.org/10.1167/15.9.2. [PubMed] [Article]

Hou, F., Zhao, Y., Lesmes, L. A., Bex, P., Yu, D., & Lu, Z.-L. (2018). Bayesian adaptive assessment of the reading function for vision: The qReading method. *Journal of Vision*, *18*(9):6, 1–15, https://doi.org/10.1167/18.9.6. [PubMed] [Article]

Huber, A. (1976). *Eye signs and symptoms in brain tumors*. St. Louis, MO: Mosby.

ICD-10 Glaucoma Staging Definitions. (2015, March 11). Retrieved from www.aao.org/practice-management/news-detail/icd-10-glaucoma-staging-definitions

James, A. C. (2003). The pattern-pulse multifocal visual evoked potential. *Investigative Ophthalmology & Visual Science*, *44*(2), 879–890.

Johnson, C. A. (2002). Recent developments in automated perimetry in glaucoma diagnosis and management. *Current Opinion in Ophthalmology*, *13*(2), 77–84.

Johnson, C. A., Chauhan, B. C., & Shapiro, L. R. (1992). Properties of staircase procedures for estimating thresholds in automated perimetry. *Investigative Ophthalmology & Visual Science*, *33*(10), 2966–2974.

Johnson, C. A., & Keltner, J. L. (1983). Incidence of visual field loss in 20,000 eyes and its relationship to driving performance. *Archives of Ophthalmology*, *101*(3), 371–375.

Johnson, C. A., & Samuels, S. J. (1997). Screening for glaucomatous visual field loss with frequency-doubling perimetry. *Investigative Ophthalmology & Visual Science*, *38*(2), 413–425.

Johnson, C. A., Wall, M., & Thompson, H. S. (2011). A history of perimetry and visual field testing. *Optometry and Vision Science*, *88*(1), E8–E15.

Kardon, R. H. (1992). Pupil perimetry. *Current Opinion in Ophthalmology*, *3*(5): 565–570.

Kardon, R. H., Kirkali, P. A., & Thompson, H. S. (1991). Automated pupil perimetry. Pupil field mapping in patients and normal subjects. *Ophthalmology*, *98*(4), 485–496.

Katz, J., Gilbert, D., Quigley, H. A., & Sommer, A. (1997). Estimating progression of visual field loss in glaucoma. *Ophthalmology*, *104*(6), 1017–1025.

Katz, J., & Sommer, A. (1987). A longitudinal study of the age-adjusted variability of automated visual fields. *Archives of Ophthalmology*, *105*(8), 1083–1086.

Keltgen, K. M., & Swanson, W. H. (2012). Estimation of spatial scale across the visual field using sinusoidal stimuli. *Investigative Ophthalmology & Visual Science*, *53*(2), 633–639.

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. *Archives of Ophthalmology*, *118*(9), 1187–1194.

Kim, W., Pitt, M. A., Lu, Z.-L., & Myung, J. I. (2017). Planning beyond the next trial in adaptive experiments: A dynamic programming approach. *Cognitive Science*, *41*(8), 2234–2252.

Kim, W., Pitt, M. A., Lu, Z.-L., Steyvers, M., &

Myung, J. I. (2014). A hierarchical adaptive approach to optimal experimental design. *Neural Computation*, 26(11), 2465–2492, https://doi.org/10.1162/NECO_a_00654.

King, A., Taguri, A., Wadood, A., & 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 for Clinical and Experimental Ophthalmology*, 240(6), 481–487, https://doi.org/10.1007/s00417-002-0482-y.

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(7), 885–912.

King-Smith, P. E., & Rose, D. (1997). Principles of an adaptive method for measuring the slope of the psychometric function. *Vision Research*, 37(12), 1595–1604.

Klein, S. A. (2001). Measuring, estimating, and understanding the psychometric function: A commentary. *Perception & Psychophysics*, 63(8), 1421–1455.

Klistorner, A. I., Graham, S. L., Grigg, J. R., & Billson, F. A. (1998). Multifocal topographic visual evoked potential: Improving objective detection of local visual field defects. *Investigative Ophthalmology & Visual Science*, 39(6), 937–950.

Kononenko, I. (2001). Machine learning for medical diagnosis: History, state of the art and perspective. *Artificial Intelligence in Medicine*, 23(1), 89–109.

Kontsevich, L. L., & Tyler, C. W. (1999). Bayesian adaptive estimation of psychometric slope and threshold. *Vision Research*, 39(16), 2729–2737.

Kujala, J. V., & Lukka, T. J. (2006). Bayesian adaptive estimation: The next dimension. *Journal of Mathematical Psychology*, 50(4), 369–389, https://doi.org/10.1016/j.jmp.2005.12.005.

Lachenmayr, B. J., Drance, S. M., Douglas, G. R., & Mikelberg, F. S. (1991). Light-sense, flicker and resolution perimetry in glaucoma: A comparative study. *Graefe's Archive for Clinical and Experimental Ophthalmology*, 229(3), 246–251.

Lachenmayr, B. J., Kojetinsky, S., Ostermaier, N., Angstwurm, K., Vivell, P. M., & Schaumberger, M. (1994). The different effects of aging on normal sensitivity in flicker and light-sense perimetry. *Investigative Ophthalmology & Visual Science*, 35(6), 2741–2748.

Landa, G., Rosen, R. B., Garcia, P. M., & Seiple, W. H. (2010). Combined three-dimensional spectral OCT/SLO topography and microperimetry: Steps toward achieving functional spectral OCT/SLO. *Ophthalmic Research*, 43(2), 92–98.

Larson, A. M., & Loschky, L. C. (2009). The contributions of central versus peripheral vision to scene gist recognition. *Journal of Vision*, 9(10):6, 1–16, https://doi.org/10.1167/9.10.6. [PubMed] [Article]

Lee, J.-W. (2019). Visual field test. In J.-S. Lee (Ed.), *Primary eye examination: A comprehensive guide to diagnosis* (pp. 179–207), Singapore: Springer Singapore, https://doi.org/10.1007/978-981-10-6940-6_14.

Leek, M. R. (2001). Adaptive procedures in psychophysical research. *Perception & Psychophysics*, 63(8), 1279–1292.

Lemmink, K. A., Dijkstra, B., & Visscher, C. (2005). Effects of limited peripheral vision on shuttle sprint performance of soccer players. *Perceptual and Motor Skills*, 100(1), 167–175.

Lennie, P. (1994). *Measurement of visual field and visual acuity for disability determination*. Washington, DC: National Academies.

Lesmes, L. A., Jeon, S.-T., Lu, Z.-L., & Dosher, B. A. (2006). Bayesian adaptive estimation of threshold versus contrast external noise functions: The quick TvC method. *Vision Research*, 46(19), 3160–3176, https://doi.org/10.1016/j.visres.2006.04.022.

Lesmes, L. A., Lu, Z.-L., Baek, J., & Albright, T. D. (2010). Bayesian adaptive estimation of the contrast sensitivity function: The quick CSF method. *Journal of Vision*, 10(3):17, 1–21, https://doi.org/10.1167/10.3.17. [PubMed] [Article]

Lesmes, L. A., Lu, Z.-L., Baek, J., Tran, N., Dosher, B. A., & Albright, T. D. (2015). Developing Bayesian adaptive methods for estimating sensitivity thresholds (d′) in Yes-No and forced-choice tasks. *Frontiers in Psychology*, 6: 1070, https://doi.org/10.3389/fpsyg.2015.01070.

Levi, D. M., & Carney, T. (2009). Crowding in peripheral vision: Why bigger is better. *Current Biology*, 19(23), 1988–1993.

Lewis, R. A., Johnson, C. A., Keltner, J. L., & Labermeier, P. K. (1986). Variability of quantitative automated perimetry in normal observers. *Ophthalmology*, 93(7), 878–881.

Lim, L. S., Mitchell, P., Seddon, J. M., Holz, F. G., & Wong, T. Y. (2012). Age-related macular degeneration. *The Lancet*, 379(9827), 1728–1738, https://doi.org/10.1016/S0140-6736(12)60282-7.

Lloyd, R., Harris, J., Wadhwa, S., & Chambers, W. (2008). Food and Drug Administration approval

process for ophthalmic drugs in the US. *Current Opinion in Ophthalmology*, *19*(3), 190–194.

Lloyd, R., Wadhwa, S., Eydelman, M., & Kramm, R. L. (2011). *The FDA's role in shaping the future of glaucoma care*. Retrieved from http://glaucomatoday.com/2011/08/the-fdas-role-in-shaping-the-future-of-glaucoma-care/

Lovie-Kitchin, J. E., Woods, R. L., Hassan, S. E., & Soong, G. P. (2001). Visual field size and reduced mobility performance. *Investigative Ophthalmology & Visual Science*, *42*, S857–S857.

Lovie-Kitchin, J. E., Soong, G. P., Hassan, S. E., & Woods, R. L. (2010). Visual field size criteria for mobility rehabilitation referral. *Optometry and Vision Science*, *87*(12), E948–E957.

Lu, Z.-L., & Dosher, B. A. (2008). Characterizing observers using external noise and observer models: Assessing internal representations with external noise. *Psychological Review*, *115*(1), 44–82.

Lu, Z.-L., & Dosher, B. A. (2013). *Visual psychophysics: From laboratory to theory*. Cambridge, MA: MIT Press.

Lu, Z.-L., Xu, P., Lesmes, L. A., & Yu, D. (n.d.). *SYSTEMS AND METHODS FOR MEASURING VISUAL FUNCTIONS*. US patent application, #62/637, 653.

Lu, Z.-L., Zhang, P., Zhao, Y., & Dosher, B. (2018). Evaluating the performance of the staircase and quick Change Detection methods in measuring perceptual learning. *Journal of Vision*, *18*(10):256, https://doi.org/10.1167/18.10.256. [Abstract]

Lu, Z.-L., Zhao, Y., Lesmes, L. A., Dorr, M., & Bex, P. (2019). Unbiased Threshold Estimates in Bayesian Adaptive qCSF and qFC with Mismatched Psychometric Function Slopes. *Investigative Ophthalmology & Visual Science*, *60*(9), 3908–3908.

Markowitz, S. N. (2006). Principles of modern low vision rehabilitation. *Canadian Journal of Ophthalmology*, *41*(3), 289–312.

Marron, J. A., & Bailey, I. L. (1982). Visual factors and orientation-mobility performance. *American Journal of Optometry and Physiological Optics*, *59*(5), 413–426.

Martin, L. (2005). Rarebit and frequency-doubling technology perimetry in children and young adults. *Acta Ophthalmologica Scandinavica*, *83*(6), 670–677.

Martin, L., & Wanger, P. (2004). New perimetric techniques: A comparison between rarebit and frequency doubling technology perimetry in normal subjects and glaucoma patients. *Journal of Glaucoma*, *13*(4), 268–272.

McKee, S. P., Klein, S. A., & Teller, D. Y. (1985). Statistical properties of forced-choice psychometric functions: Implications of probit analysis. *Perception & Psychophysics*, *37*(4), 286–298.

McKendrick, A. M. (2005). Recent developments in perimetry: Test stimuli and procedures. *Clinical and Experimental Optometry*, *88*(2), 73–80, https://doi.org/10.1111/j.1444-0938.2005.tb06671.x.

Menke, M. N., Sato, E., Van De Velde, F. J., & Feke, G. T. (2006). Combined use of SLO microperimetry and OCT for retinal functional and structural testing. *Graefe's Archive for Clinical and Experimental Ophthalmology*, *244*(5), 634–638.

Mills, R. P., Budenz, D. L., Lee, P. P., Noecker, R. J., Walt, J. G., Siegartel, L. R., . . . Doyle, J. J. (2006). Categorizing the stage of glaucoma from pre-diagnosis to end-stage disease. *American Journal of Ophthalmology*, *141*(1), 24–30.

Morales, J., Weitzman, M. L., & González de la Rosa, M. (2000). Comparison between tendency-oriented perimetry (TOP) and Octopus threshold perimetry. *Ophthalmology*, *107*(1), 134–142, https://doi.org/10.1016/S0161-6420(99)00026-3.

Murray, I. C., Fleck, B. W., Brash, H. M., MacRae, M. E., Tan, L. L., & Minns, R. A. (2009). Feasibility of saccadic vector optokinetic perimetry: A method of automated static perimetry for children using eye tracking. *Ophthalmology*, *116*(10), 2017–2026.

Murray, I., Perperidis, A., Brash, H., Cameron, L., McTrusty, A., Fleck, B., & Minns, R. (2013). Saccadic Vector Optokinetic Perimetry (SVOP): A novel technique for automated static perimetry in children using eye tracking. *35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society* (pp. 3186–3189). New York, NY: IEEE.

National Research Council. (2002). *Visual impairments: Determining eligibility for social security benefits*. Washington (DC): National Academies Press.

Ng, M., Sample, P. A., Pascual, J. P., Zangwill, L. M., Girkin, C. A., Liebmann, J. M., . . . Racette, L. (2012). Comparison of visual field severity classification systems for glaucoma. *Journal of Glaucoma*, *21*(8), 551–561.

Nguyen, D. T., Fahimi, A., Fink, W., Nazemi, P. P., Kim, J. K., & Sadun, A. A. (2009). Novel 3D computer-automated threshold Amsler grid visual field testing of scotomas in patients with glaucoma. *European Journal of Ophthalmology*, *19*(5), 776–782.

O'Brien, C., Poinoosawmy, D., Wu, J., & Hitchings, R. (1994). Evaluation of the Humphrey FASTPAC threshold program in glaucoma. *British Journal of*

Ophthalmology, *78*(7), 516–519, https://doi.org/10.1136/bjo.78.7.516.

Okada, K., Yamamoto, S., Mizunoya, S., Hoshino, A., Arai, M., & Takatsuna, Y. (2006). Correlation of retinal sensitivity measured with fundus-related microperimetry to visual acuity and retinal thickness in eyes with diabetic macular edema. *Eye*, *20*(7), 805–809.

Okamoto, Y., Okamoto, F., Hiraoka, T., Yamada, S., & Oshika, T. (2008). Vision-related quality of life in patients with pituitary adenoma. *American Journal of Ophthalmology*, *146*(2), 318–322.

Oshika, T., Klyce, S. D., Applegate, R. A., & Howland, H. C. (1999). Changes in corneal wavefront aberrations with aging. *Investigative Ophthalmology & Visual Science*, *40*(7), 1351–1355.

Oshika, T., Okamoto, C., Samejima, T., Tokunaga, T., & Miyata, K. (2006). Contrast sensitivity function and ocular higher-order wavefront aberrations in normal human eyes. *Ophthalmology*, *113*(10), 1807–1812.

Palmer, J., Ames, C. T., & Lindsey, D. T. (1993). Measuring the effect of attention on simple visual search. *Journal of Experimental Psychology: Human Perception and Performance*, *19*(1), 108–130.

Palmer, J., Verghese, P., & Pavel, M. (2000). The psychophysics of visual search. *Vision Research*, *40*(10–12), 1227–1268.

Papageorgiou, E., Hardiess, G., Schaeffel, F., Wiethoelter, H., Karnath, H.-O., Mallot, H., . . . Schiefer, U. (2007). Assessment of vision-related quality of life in patients with homonymous visual field defects. *Graefe's Archive for Clinical and Experimental Ophthalmology*, *245*(12), 1749–1758, https://doi.org/10.1007/s00417-007-0644-z.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, *10*(4), 437–442.

Phu, J., Khuu, S. K., Zangerl, B., & Kalloniatis, M. (2017). A comparison of Goldmann III, V and spatially equated test stimuli in visual field testing: the importance of complete and partial spatial summation. *Ophthalmic & Physiological Optics*, *37*(2), 160–176. https://doi.org/10.1111/opo.12355

Qiu, Z., Xu, P., Zhou, Y., & Lu, Z.-L. (2007). Spatial vision deficit underlies poor sine-wave motion direction discrimination in anisometropic amblyopia. *Journal of Vision*, *7*(11):7, 1–15, https://doi.org/10.1167/7.11.7. [PubMed] [Article]

Quigley, H. A., Tielsch, J. M., Katz, J., & Sommer, A. (1996). Rate of progression in open-angle glaucoma estimated from cross-sectional prevalence of visual field damage. *American Journal of Ophthalmology*, *122*(3), 355–363.

Rajan, M. S., Bremner, F. D., & Riordan-Eva, P. (2002). Pupil perimetry in the diagnosis of functional visual field loss. *Journal of the Royal Society of Medicine*, *95*(10), 498–500.

Ramulu, P. (2009). Glaucoma and disability: Which tasks are affected, and at what stage of disease? *Current Opinion in Ophthalmology*, *20*(2), 92–98.

Ramulu, P. Y., West, S. K., Munoz, B., Jampel, H. D., & Friedman, D. S. (2009). Glaucoma and reading speed: The Salisbury Eye Evaluation project. *Archives of Ophthalmology*, *127*(1), 82–87.

Rogers, T. J., & Landers, D. M. (2005). Mediating effects of peripheral vision in the life event stress/athletic injury relationship. *Journal of Sport and Exercise Psychology*, *27*(3), 271–288.

Rowe, F. J., Cheyne, C. P., García-Fiñana, M., Noonan, C. P., Howard, C., Smith, J., & Adeoye, J. (2015). Detection of visual field loss in pituitary disease: Peripheral kinetic versus central static. *Neuro-Ophthalmology*, *39*(3), 116–124, https://doi.org/10.3109/01658107.2014.990985.

Sample, P. A., & Weinreb, R. N. (1990). Color perimetry for assessment of primary open-angle glaucoma. *Investigative Ophthalmology & Visual Science*, *31*(9), 1869–1875.

Sample, P. A., & Weinreb, R. N. (1992). Progressive color visual field loss in glaucoma. *Investigative Ophthalmology & Visual Science*, *33*(6), 2068–2071.

Schiefer, U., Pascual, J. P., Edmunds, B., Feudner, E., Hoffmann, E. M., Johnson, C. A., . . . Paetzold, J. (2009). Comparison of the New Perimetric GATE Strategy with Conventional Full-Threshold and SITA Standard Strategies. Investigative Opthalmology & Visual Science, *50*(1), 488. https://doi.org/10.1167/iovs.08-2229

Schulzer, M., Airaksinen, P. J., Alward, W. L., Amyot, M., Anderson, D. R., Balazsi, G., . . . Desjardins, D. (1994). Errors in the diagnosis of visual field progression in normal-tension glaucoma. *Ophthalmology*, *101*(9), 1589–1595.

Sharma, P., Sample, P. A., Zangwill, L. M., & Schuman, J. S. (2008). Diagnostic tools for glaucoma detection and management. *Survey of Ophthalmology*, *53*(6), S17–S32.

Shaw, M. L. (1980). Identifying attentional and decision-making components in information processing. *Attention and performance VIII* (pp. 277–295). Hillsdale, NJ: Lawrence Erlbaum Associates.

Shepard, T. G., Hou, F., Bex, P. J., Lesmes, L. A., Lu, Z.-L., & Yu, D. (2019). Assessing reading perfor-

mance in the periphery with a Bayesian adaptive approach: The qReading method. *Journal of Vision*, 19(5):5, 1–14, https://doi.org/10.1167/19.5.5. [PubMed] [Article]

Simpson, W. A. (1989). The step method: A new adaptive psychophysical procedure. *Perception & Psychophysics*, 45(6), 572–576.

Smith, S. D., Katz, J., & Quigley, H. A. (1996). Analysis of progressive change in automated visual fields in glaucoma. *Investigative Ophthalmology & Visual Science*, 37(7), 1419–1428.

Smythies, J. (1996). A note on the concept of the visual field in neurology, psychology, and visual neuroscience. *Perception*, 25(3), 369–371.

Sperling, G. & Lu, Z.-L. (1998). A systems analysis of visual motion perception. In T. Watanabe (Ed.), *High-level motion processing: Computational, neurobiological, and psychophysical perspectives* (pp. 154–183). Cambridge, MA: MIT Press.

Stewart, W. C., & Hunt, H. H. (1993). Threshold variation in automated perimetry. *Survey of Ophthalmology*, 37(5), 353–361.

Strasburger, H., Rentschler, I., & Jüttner, M. (2011). Peripheral vision and pattern recognition: A review. *Journal of Vision*, 11(5):13, 1–82, https://doi.org/10.1167/11.5.13. [PubMed] [Article]

Susanna, R., Jr., & Vessani, R. M. (2009). Staging glaucoma patient: Why and how? *The Open Ophthalmology Journal*, 3, 59–64.

Swanson, W. H., Horner, D. G., Dul, M. W., & Malinovsky, V. E. (2014). Choice of stimulus range and size can reduce test-retest variability in glaucomatous visual field defects. *Translational Vision Science & Technology*, 3(5): 6.

Swanson, W. H., Malinovsky, V. E., Dul, M. W., Malik, R., Torbit, J. K., Sutton, B. M., & Horner, D. G. (2014). Contrast sensitivity perimetry and clinical measures of glaucomatous damage. *Optometry and Vision Science*, 91(11), 1302–1311.

Tanner, T. G. (2008). Generalized adaptive procedure for psychometric measurement. *31st European Conference on Visual Perception. 37 (ECVP Abstract Supplement)*, 93–93.

Thompson, H. S., Montague, P., Cox, T. A., & Corbett, J. J. (1982). The relationship between visual acuity, pupillary defect, and visual field loss. *American Journal of Ophthalmology*, 93(6), 681–688.

Thornton, C., Livermore, G., Stapleton, D., Kregel, J., Silva, T., O'Day, B., . . . Edwards, M. (2004). *Evaluation of the Ticket to Work program: Initial evaluation report*. Washington, DC: Mathematica Policy Research.

Townend, B. S., Sturm, J. W., Petsoglou, C., O'Leary, B., Whyte, S., & Crimmins, D. (2007). Perimetric homonymous visual field loss post-stroke. *Journal of Clinical Neuroscience*, 14(8), 754–756.

Treutwein, B. (1995). Adaptive psychophysical procedures. *Vision Research*, 35(17), 2503–2522.

van Gaalen, K. W., Jansonius, N. M., Koopmans, S. A., Terwee, T., & Kooijman, A. C. (2009). Relationship between contrast sensitivity and spherical aberration: Comparison of 7 contrast sensitivity tests with natural and artificial pupils in healthy eyes. *Journal of Cataract & Refractive Surgery*, 35(1), 47–56.

Wall, M., Brito, C. F., & Kutzko, K. (1997). Motion detection perimetry properties and results. *Perimetry Update*, 96/97, 21–33.

Wall, M., & Ketoff, K. M. (1995). Random dot motion perimetry in patients with glaucoma and in normal subjects. *American Journal of Ophthalmology*, 120(5), 587–596.

Wall, M., & May, D. R. (1987). Threshold Amsler grid testing in maculopathies. *Ophthalmology*, 94(9), 1126–1133.

Wall, M., Woodward, K. R., Doyle, C. K., & Artes, P. H. (2009). Repeatability of automated perimetry: A comparison between standard automated perimetry with stimulus size III and V, matrix, and motion perimetry. *Investigative Ophthalmology & Visual Science*, 50(2), 974–979.

Wang, F., Javitt, J. C., Rowe, M., & Meng, K. (1996). Measuring the impact of glaucoma and its treatment on quality of life: The glaucoma disability index. *Investigative Ophthalmology and Visual Science*, 37(3): 0146–0404.

Watson, A. B. (2017). QUEST+: A general multidimensional Bayesian adaptive psychometric method. *Journal of Vision*, 17(3):10, 1–27, https://doi.org/10.1167/17.3.10. [PubMed] [Article]

Watson, A. B., & Pelli, D. G. (1983). QUEST: A Bayesian adaptive psychometric method. *Perception & Psychophysics*, 33(2), 113–120.

Weijland, A., Fankhauser, F., Bebie, H., & Flammer, J. (2004). *Automated perimetry: Visual field digest*. Köniz, Switzerland: Haag-Streit AG.

Weinreb, R. N., & Kaufman, P. L. (2009). The glaucoma research community and FDA look to the future: A report from the NEI/FDA CDER Glaucoma Clinical Trial Design and Endpoints Symposium. *Investigative Ophthalmology & Visual Science*, 50(4), 1497–1505.

Weinreb, R. N., & Kaufman, P. L. (2011). Glaucoma research community and FDA look to the future, II: NEI/FDA Glaucoma Clinical Trial Design and Endpoints Symposium: Measures of structural change and visual function. *Investigative Ophthalmology & Visual Science*, *52*(11), 7842–7851.

Wen, J. C., Lee, C. S., Keane, P. A., Xiao, S., Rokem, A. S., Chen, P. P., . . . Lee, A. Y. (2019). Forecasting future Humphrey Visual Fields using deep learning. *PLoS One*, *14*(4), e0214875, https://doi.org/10.1371/journal.pone.0214875.

Werner, E. B., Saheb, N., & Thomas, D. (1982). Variability of static visual threshold responses in patients with elevated IOPs. *Archives of Ophthalmology*, *100*(10), 1627–1631.

Westheimer, G. (1965). Visual acuity. *Annual Review of Psychology*, *16*(1), 359–380.

Wichmann, F. A., & Hill, N. J. (2001). The psychometric function: I. Fitting, sampling, and goodness of fit. *Perception & Psychophysics*, *63*(8), 1293–1313.

Wilensky, J. T., & Joondeph, B. C. (1984). Variation in visual field measurements with an automated perimeter. *American Journal of Ophthalmology*, *97*(3), 328–331, https://doi.org/10.1016/0002-9394(84)90631-7.

Wixon, B., & Strand, A. (2013). Identifying SSA's Sequential Disability Determination Steps Using Administrative Data (SSRN Scholarly Paper No. ID 2277703). Retrieved from Social Science Research Network website: https://papers.ssrn.com/abstract=2277703.

Wood, J. M. (2002). Age and visual impairment decrease driving performance as measured on a closed-road circuit. *Human Factors*, *44*(3), 482–494.

Xu, P., Lu, Z.-L., Qiu, Z., & Zhou, Y. (2006). Identify mechanisms of amblyopia in Gabor orientation identification with external noise. *Vision Research*, *46*(21), 3748–3760.

Xu, P., Lu, Z.-L., Wang, X., Dosher, B. A., Zhou, J., Zhang, D., & Zhou, Y. (2010). Category and perceptual learning in subjects with treated Wilson's disease. *PLoS One*, *5*(3), e9635.

Yoshiyama, K. K., & Johnson, C. A. (1997). Which method of flicker perimetry is most effective for detection of glaucomatous visual field loss? *Investigative Ophthalmology & Visual Science*, *38*(11), 2270–2277.

Yu, D., Cheung, S.-H., Legge, G. E., & Chung, S. T. (2010). Reading speed in the peripheral visual field of older adults: Does it benefit from perceptual learning? *Vision Research*, *50*(9), 860–869.

Zhao, Y., Lesmes, L., & Lu, Z.-L. (2019). Efficient assessment of the time course of perceptual sensitivity change. *Vision Research*, *154*, 21–43, https://doi.org/10.1016/j.visres.2018.10.009.

Appendix A: Additional experimental results (Subjects 2–6)

The estimated light-sensitivity visual-field maps of the left and right eyes from Subjects 2 through 6 are shown in Figures A1 through A5.

[OS]   [OD]

Estimated VFM (qVFM 300 trial) | Estimated VFM (qYN 1200 trial) | Estimated VFM (qVFM 300 trial) | Estimated VFM (qYN 1200 trial)

Estimated VFM (qVFM 300 trial) — OS

19.5	20.9	21.1	21.3	20.8	21.3	21.7	21.2	20.8	20.7
21.7	21.6	21.8	22.0	22.1	22.4	22.4	22.3	22.1	21.6
22.3	22.3	22.2	22.7	22.7	23.1	23.0	22.8	22.8	22.8
22.5	22.9	23.1	23.2	23.3	23.4	23.4	23.5	23.3	23.1
22.7	22.9	23.1	23.5	23.6	23.7	23.7	23.6	23.3	23.3
22.8	23.1	10.5	23.5	23.7	23.8	23.7	23.6	23.3	23.3
23.1	23.0	23.5	23.6	23.6	23.6	23.6	23.4	23.3	23.1
22.5	22.6	23.1	23.5	23.4	23.3	23.4	23.1	23.1	23.2
22.8	22.6	22.8	23.1	23.0	22.9	22.9	23.1	22.9	22.7
21.9	22.3	22.1	22.6	22.6	22.6	22.3	22.3	22.7	22.3

Estimated VFM (qYN 1200 trial) — OS

18.0	20.8	21.3	21.4	21.7	21.8	22.6	21.3	21.1	19.8
21.9	22.4	21.9	21.7	21.9	22.4	22.7	21.1	21.9	19.9
22.8	22.8	21.6	23.2	22.4	22.9	23.1	22.6	21.5	23.2
22.6	22.7	23.0	23.1	23.7	23.7	23.5	23.7	23.2	18.6
22.9	22.2	22.3	23.5	24.1	24.1	23.9	23.7	22.9	23.4
22.9	22.7	10.4	23.6	24.1	24.0	23.5	23.7	23.5	23.3
22.8	22.4	23.3	23.9	23.7	23.5	23.6	23.7	23.0	23.1
21.6	22.2	22.5	23.6	23.3	23.4	23.1	22.9	22.8	23.4
23.4	22.7	22.7	22.6	22.4	22.4	22.3	22.7	22.8	22.5
10.2	19.2	20.6	22.3	21.3	22.1	21.4	21.5	22.1	21.9

Estimated VFM (qVFM 300 trial) — OD

21.1	20.9	21.0	21.4	21.1	21.1	20.9	20.4	20.5	20.2
21.9	22.1	21.9	21.8	21.7	21.1	21.6	20.9	21.1	20.7
22.6	22.4	21.9	22.3	22.3	22.2	22.0	21.5	21.3	20.8
22.8	22.5	22.6	22.8	22.7	22.4	22.4	21.9	21.6	21.6
23.0	22.9	22.8	23.0	22.9	22.9	22.5	17.7	21.2	21.5
22.9	22.4	22.8	22.7	22.9	22.8	22.6	13.9	21.5	21.0
22.7	22.5	22.7	22.8	22.7	22.4	22.1	21.9	20.5	20.0
22.3	22.6	22.4	22.3	22.3	22.0	22.2	21.5	20.8	20.7
22.1	22.3	22.0	21.5	21.7	21.6	21.4	21.0	21.1	20.7
21.7	21.7	21.2	21.4	20.9	20.6	20.7	20.5	20.5	19.7

Estimated VFM (qYN 1200 trial) — OD

20.5	20.9	18.9	21.7	20.8	20.5	21.4	21.1	21.6	21.3
21.3	21.5	21.1	21.1	21.1	21.4	21.1	20.7	21.0	20.7
22.5	21.8	22.1	21.7	21.9	21.8	22.2	20.7	20.9	20.9
22.9	22.2	22.5	22.9	22.3	22.5	21.5	21.0	20.2	20.8
23.3	22.1	21.7	23.1	23.6	23.8	22.3	15.5	21.2	21.3
22.9	22.0	22.6	22.5	23.4	23.7	22.8	9.9	21.3	20.3
22.6	22.1	23.1	23.3	22.8	22.5	21.6	22.3	20.0	19.7
22.7	22.6	22.7	23.0	22.7	22.2	22.0	21.5	20.4	20.2
22.7	22.9	21.6	21.3	20.6	21.8	21.5	21.0	21.0	21.8
23.2	22.3	21.3	21.7	20.5	20.9	21.1	20.9	21.1	21.1

HWCI (qVFM 300 trial) — OS

.61	.46	.49	.45	.48	.45	.42	.49	.49	.47
.45	.40	.38	.36	.37	.30	.33	.34	.35	.41
.37	.33	.34	.30	.30	.25	.26	.29	.29	.29
.35	.30	.26	.26	.23	.22	.22	.21	.26	.26
.28	.28	.23	.21	.16	.17	.17	.20	.24	.23
.29	.24	.92	.22	.15	.16	.20	.23	.24	.25
.29	.28	.22	.19	.18	.21	.20	.22	.26	.26
.30	.29	.26	.22	.20	.24	.23	.27	.26	.24
.32	.30	.29	.27	.28	.28	.26	.27	.29	.30
.40	.34	.37	.31	.32	.31	.33	.33	.31	.34

HWCI (qYN 1200 trial) — OS

.78	.54	.42	.47	.40	.37	.30	.47	.42	.54
.37	.37	.37	.44	.37	.33	.34	.50	.42	.54
.28	.28	.40	.30	.33	.28	.26	.30	.40	.30
.37	.30	.26	.26	.21	.20	.23	.21	.30	.65
.28	.35	.39	.26	.15	.15	.21	.20	.28	.23
.32	.34	.78	.24	.19	.16	.21	.21	.21	.25
.28	.33	.28	.18	.24	.23	.21	.20	.32	.26
.44	.39	.30	.21	.25	.23	.26	.28	.28	.28
.23	.30	.30	.37	.33	.33	.39	.30	.34	.37
.78	.69	.48	.33	.42	.35	.47	.40	.35	.37

HWCI (qVFM 300 trial) — OD

.46	.44	.42	.40	.46	.40	.45	.49	.49	.55
.35	.35	.35	.36	.35	.41	.38	.47	.42	.50
.29	.33	.37	.32	.32	.33	.33	.40	.38	.45
.28	.29	.27	.29	.26	.29	.28	.33	.38	.39
.23	.29	.27	.26	.25	.28	.30	.65	.43	.38
.25	.31	.29	.30	.27	.28	.31	.65	.37	.47
.29	.29	.29	.27	.31	.31	.30	.34	.50	.56
.34	.30	.28	.32	.31	.33	.31	.37	.48	.44
.38	.35	.35	.36	.39	.40	.38	.44	.44	.50
.43	.43	.42	.42	.44	.42	.45	.49	.49	.59

HWCI (qYN 1200 trial) — OD

.48	.45	.65	.44	.54	.48	.33	.50	.40	.47
.42	.40	.50	.42	.42	.47	.42	.54	.50	.54
.37	.37	.35	.37	.37	.37	.35	.54	.45	.45
.28	.35	.37	.28	.39	.33	.40	.45	.51	.45
.28	.39	.44	.26	.21	.20	.33	.65	.42	.42
.28	.35	.30	.37	.28	.20	.34	.65	.42	.51
.30	.35	.26	.25	.28	.37	.40	.33	.51	.54
.30	.30	.30	.32	.30	.35	.42	.40	.57	.61
.30	.28	.40	.42	.48	.37	.40	.45	.45	.37
.25	.33	.42	.44	.48	.45	.42	.45	.42	.50

SD (qVFM 300 trial) — OS

2.0	.28	.45	.80	.87	.68	.67	1.2	1.4	.93
.64	.42	.71	.30	.19	.46	.51	.65	.81	.54
.35	.46	.32	.22	.19	.29	.17	.35	.45	.54
.35	.24	.36	.13	.16	.09	.13	.26	.17	.31
.19	.12	.25	.14	.18	.17	.07	.14	.13	.23
.18	.18	.21	.16	.12	.11	.14	.14	.09	.26
.13	.18	.16	.17	.09	.15	.22	.17	.27	.09
.29	.37	.25	.15	.23	.25	.12	.31	.08	.26
.32	.14	.32	.16	.10	.53	.28	.20	.74	
.73	.24	.40	.28	.39	.34	.61	.31	.22	.68

RMSE — OS

2.3	.26	.42	.69	1.1	.77	1.0	1.0	1.2	1.2
.59	.88	.63	.43	.24	.40	.54	1.3	.74	1.8
.59	.61	.67	.48	.41	.29	.17	.36	1.3	.58
.30	.25	.33	.16	.37	.37	.13	.28	.20	4.5
.27	.75	.83	.13	.49	.41	.27	.21	.42	.26
.24	.36	.20	.17	.36	.22	.20	.16	.25	.23
.26	.59	.22	.31	.09	.17	.19	.31	.36	.09
.89	.45	.63	.20	.20	.25	.38	.31	.31	.30
.65	.13	.31	.52	.67	.43	.81	.45	.23	.68
11.6	3.1	1.5	.37	1.3	.66	1.0	.84	.65	.73

SD (qVFM 300 trial) — OD

.42	.45	.24	.21	.29	.30	.27	.22	.39	.48
.28	.23	.20	.34	.24	.63	.19	.11	.27	.54
.32	.25	.34	.12	.11	.27	.26	.40	.23	.55
.26	.06	.14	.12	.15	.17	.24	.31	.32	.41
.08	.23	.36	.17	.11	.17	.18	.85	.60	.43
.10	.23	.29	.21	.14	.17	.14	1.7	.24	.56
.42	.34	.18	.30	.06	.22	.41	.40	.39	.58
.22	.23	.32	.30	.26	.24	.23	.33	.64	.23
.29	.31	.25	.33	.53	.59	.13	.29	.22	.75
.28	.30	.60	.15	.55	.72	.62	.34	.53	.61

RMSE — OD

.72	.39	2.1	.39	.37	.61	.52	.72	1.1	1.1
.61	.65	.83	.72	.61	.62	.47	.16	.25	.48
.33	.58	.35	.62	.40	.48	.28	.84	.53	.48
.26	.32	.16	.12	.43	.15	.91	.99	1.4	.81
.33	.73	1.1	.19	.66	.94	.24	2.2	.52	.41
.12	.40	.32	.30	.52	.89	.25	4.2	.24	.86
.38	.52	.47	.52	.15	.24	.63	.57	.64	.59
.40	.21	.38	.68	.43	.26	.30	.29	.65	.61
.71	.73	.47	.35	1.1	.55	.14	.25	.19	1.2
1.5	.63	.53	.34	.61	.67	.71	.50	.72	1.4

X(degree)   X(degree)   X(degree)   X(degree)

Figure A1. Experimental results for Subject 2 (both left and right eyes). Estimated visual-field maps (VFMs) are presented in the first row with achromatic color maps and the second row with numerical values (unit: dB). For each visual-field location of the estimated VFM, the 68.2% half width of the credible interval is presented in the third row, and the standard deviations from four sessions of the qVFM method and the RMSE[eyes] between the qVFM and qYN methods are presented in the fourth row. The results obtained from the left eye are displayed in the first and second columns, and from the right eye in the third and fourth columns. The results from the qVFM and qYN methods are displayed in different columns.
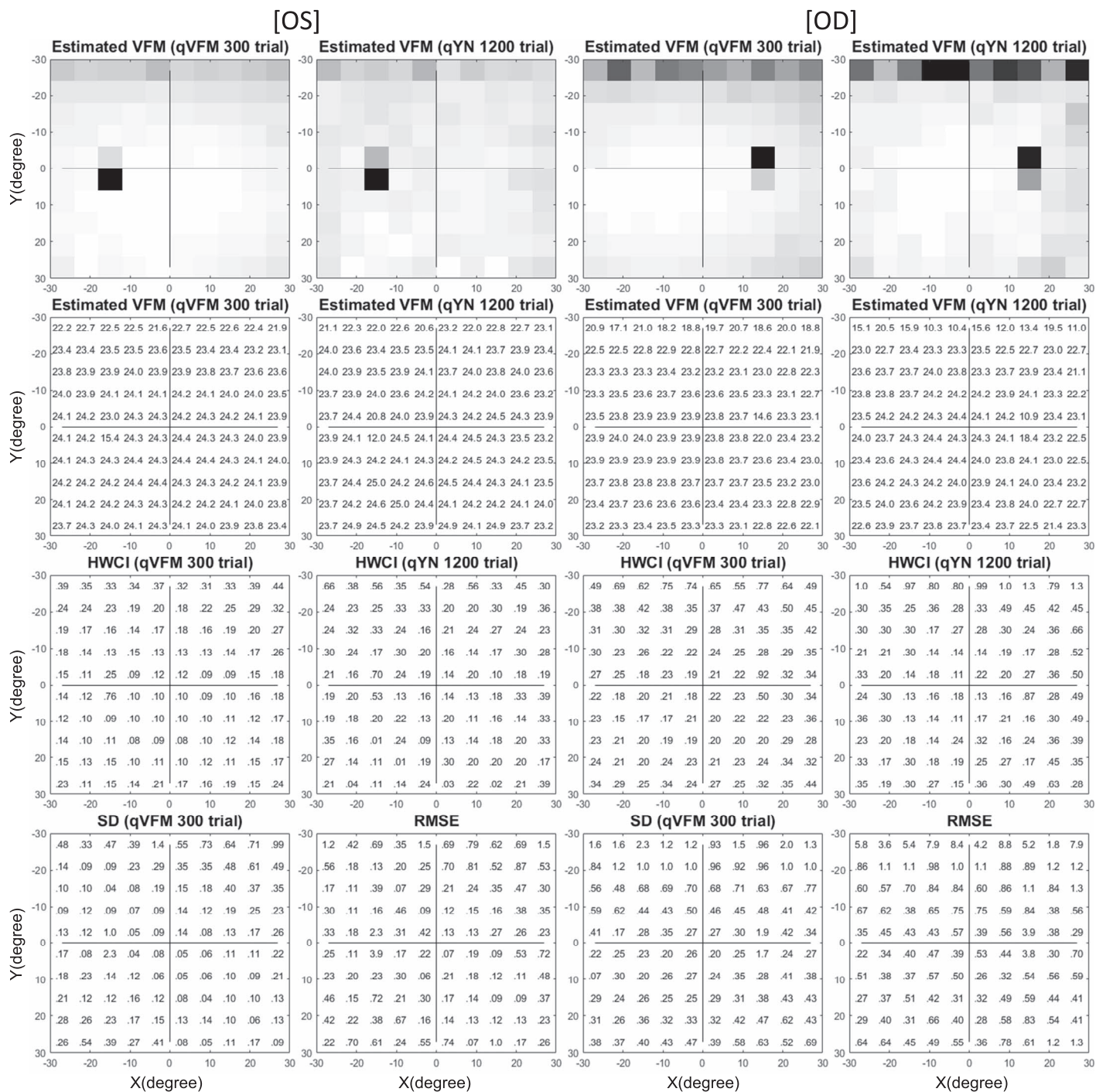
Figure A2. Experimental results for Subject 3. Details are the same as in Figure A1.
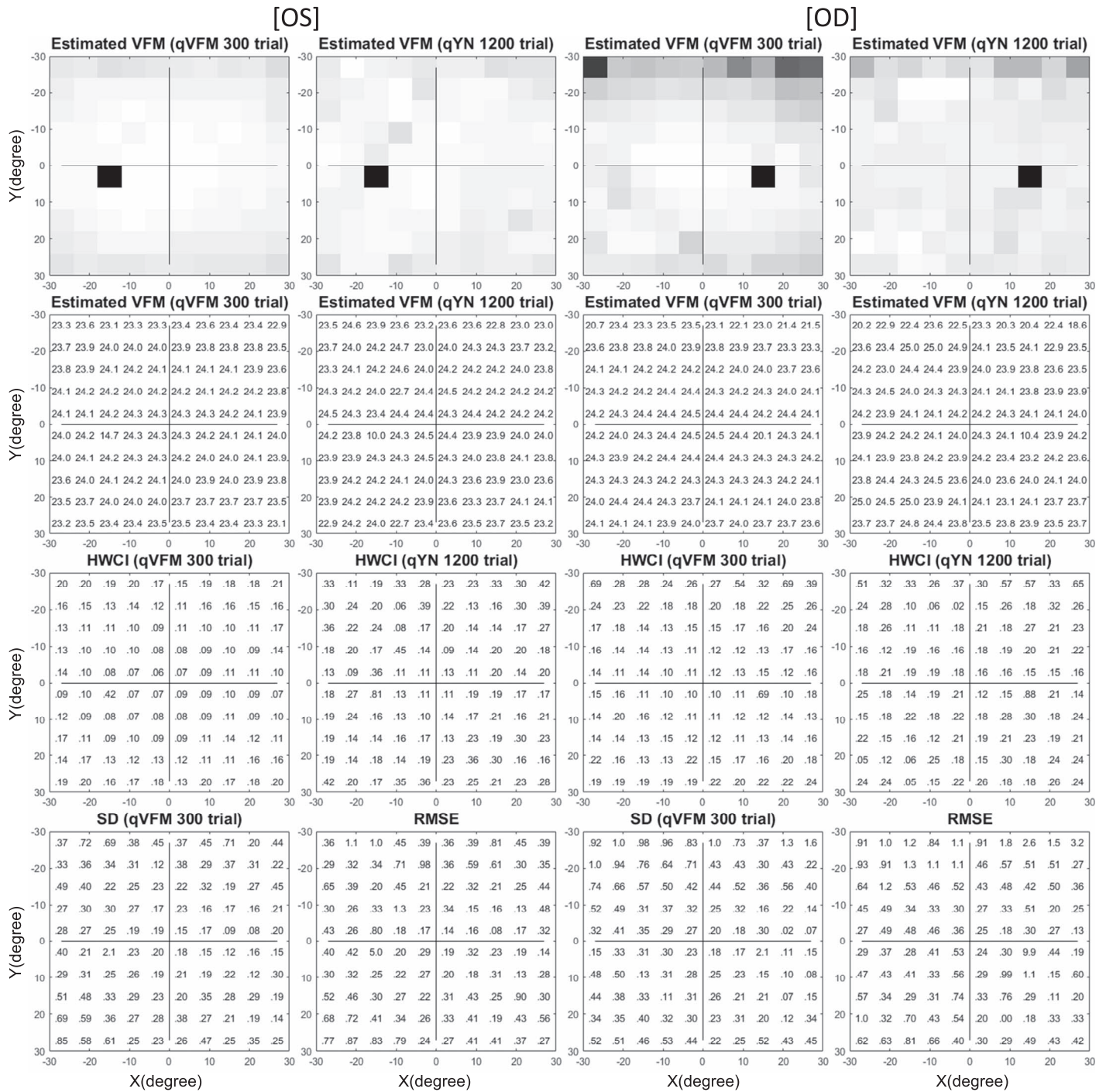
Figure A3. Experimental results for Subject 4. Details are the same as in Figure A1
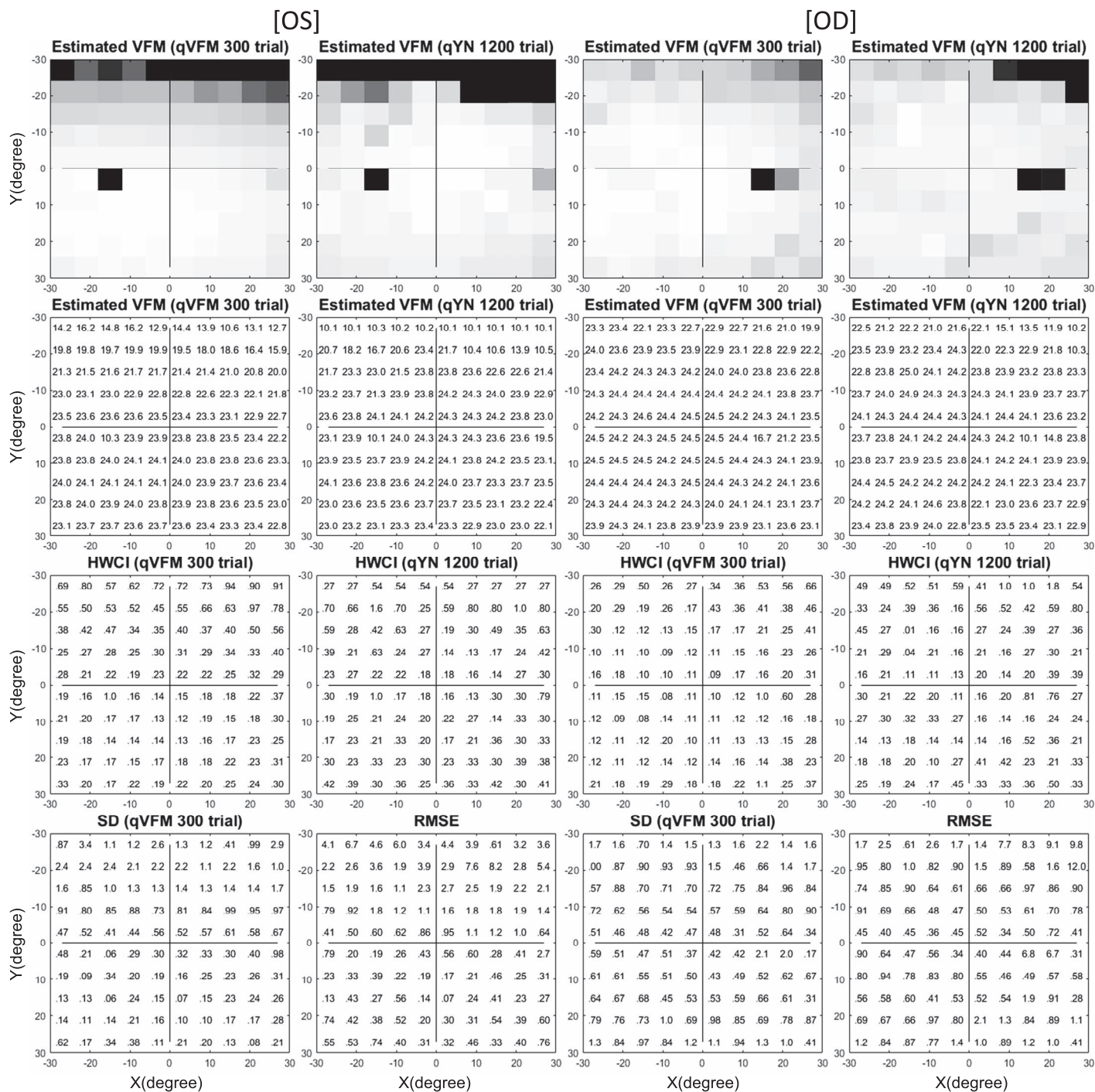
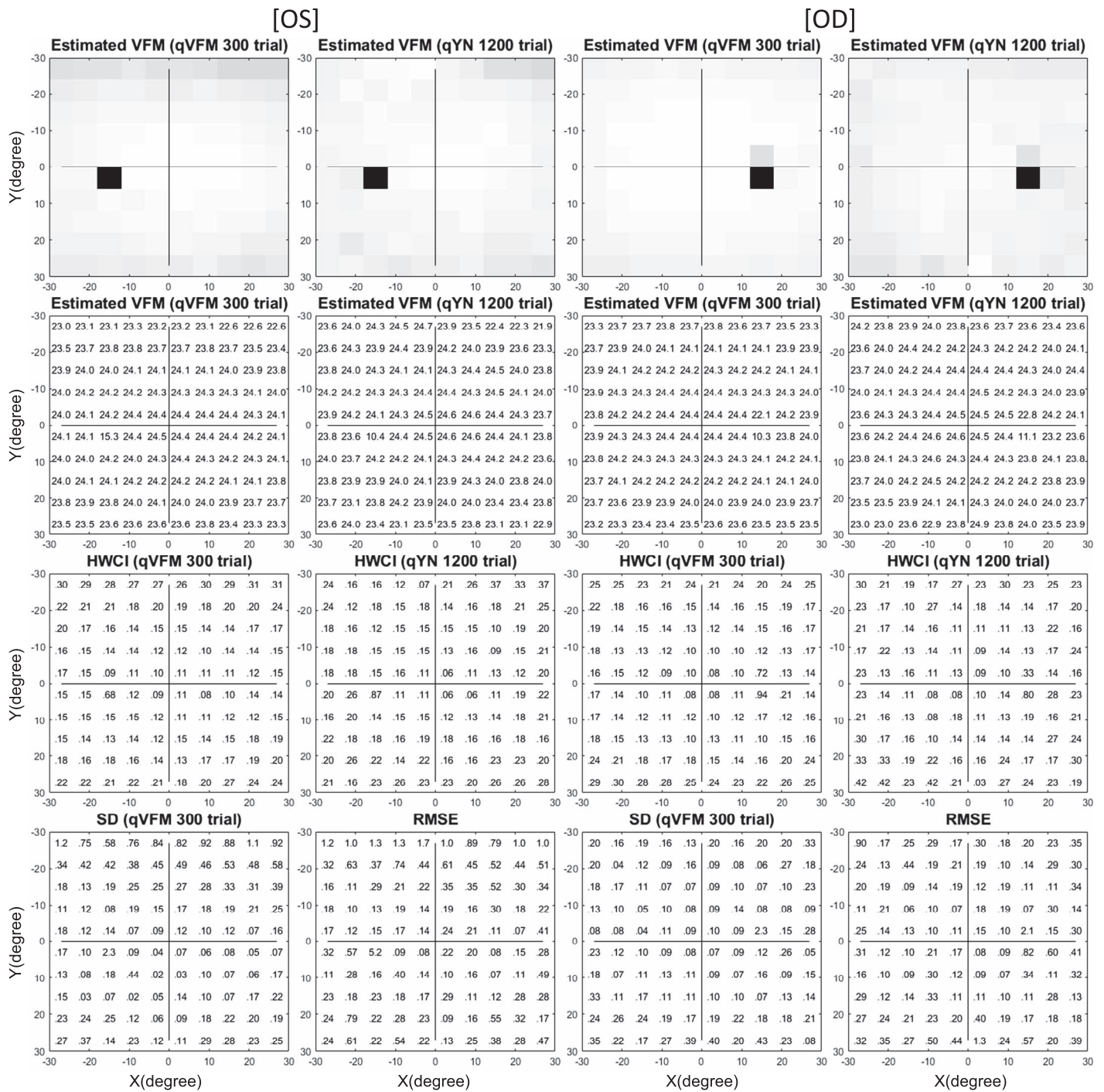Figure A4. Experimental results for Subject 5. Details are the same as in Figure A1.

Figure A5. Experimental results for Subject 6. Details are the same as in Figure A1.