OPEN ACCESS

Comparing the usefulness of a new algorithm to measure visual field using the variational Bayes linear regression in glaucoma patients, in comparison to the Swedish interactive thresholding algorithm

Hiroshi Murata,[1] Ryo Asaoka ,[1,2,3] Yuri Fujino ,[1,2,4] Masato Matsuura ,[1,5] Kazunori Hirasawa ,[5] Satoshi Shimada,[6] Nobuyuki Shoji[5]

[1]Department of Ophthalmology, The University of Tokyo, Tokyo, Japan
[2]Department of Ophthalmology, Seirei Hamamatsu General Hospital, Hamamatsu, Japan
[3]Seirei Christopher University, Hamamatsu, Japan
[4]Department of Ophthalmology, Shimane University Faculty of Medicine, Izumo, Japan
[5]Department of Ophthalmology, Kitasato University Graduate School of Medical Sciences, Kanagawa, Japan
[6]Kowa Company, Ltd, Nagoya, Japan

Correspondence to
Dr Ryo Asaoka, Department of Ophthalmology, Tokyo University, Bunkyo-ku 112-0001, Japan; ryoasa0120@mac.com

Received 31 October 2020
Revised 7 December 2020
Accepted 19 December 2020
Published Online First 13 January 2021

To cite: Murata H, Asaoka R, Fujino Y, *et al*. *Br J Ophthalmol* 2022;**106**:660–666.

ABSTRACT

Background/aims We previously reported that the visual field (VF) prediction model using the variational Bayes linear regression (VBLR) is useful for accurately predicting VF progression in glaucoma (Invest Ophthalmol Vis Sci. 2014, 2018). We constructed a VF measurement algorithm using VBLR, and the purpose of this study was to investigate its usefulness.

Method 122 eyes of 73 patients with open-angle glaucoma were included in the current study. VF measurement was performed using the currently proposed VBLR programme with AP-7700 perimetry (KOWA). VF measurements were also conducted using the Swedish interactive thresholding algorithm (SITA) standard programme with Humphrey field analyser. VF measurements were performed using the 24–2 test grid. Visual sensitivities, test–retest reproducibility and measurement duration were compared between the two algorithms.

Result Mean mean deviation (MD) values with SITA standard were −7.9 and −8.7 dB (first and second measurements), whereas those with VBLR-VF were −8.2 and −8.0 dB, respectively. There were no significant differences across these values. The correlation coefficient of MD values between the 2 algorithms was 0.97 or 0.98. Test–retest reproducibility did not differ between the two algorithms. Mean measurement duration with SITA standard was 6 min and 02 s or 6 min and 00 s (first or second measurement), whereas a significantly shorter duration was associated with VBLR-VF (5 min and 23 s or 5 min and 30 s).

Conclusion VBLR-VF reduced test duration while maintaining the same accuracy as the SITA-standard.

INTRODUCTION

Visual field (VF) testing is undoubtedly essential for the management of glaucoma. The Swedish interactive thresholding algorithm (SITA) standard programme[1] in the Humphrey field analyser (HFA, Carl Zeiss Meditec, Dublin, USA) has been widely used to measure VF since the 1980s. The main advancement compared with the former Full Threshold programme is the reduction of test duration without losing measurement quality by continuously updating the iterative maximum posterior probability estimation of the threshold using a Bayesian prior model and interrupting testing at each tested location at predetermined levels of test certainty. As a result, it was reported that the measurement duration was by up to 50% shorter[2] and test–retest reproducibility was diminished with the SITA standard programme, probably due to the suppression of patient fatigue.[3] SITA fast is a further rapid algorithm that presents the starting stimuli closer to the expected thresholds and interrupts stimulus staircases at an earlier stage by increasing error-related factor cut-off.[4] However, this allows for accepting poorer accuracy of test results and indeed, a previous study suggested that SITA fast had better test–retest reproducibility[3] and poorer measurement accuracy compared with the SIAT standard.[5] SITA faster is an even newer algorithm that enables fastest VF measurement in the SITA family by adjusting the initial intensity for age, omitting some reversals in thresholding, not providing false-negative (FN) results and also eliminating blind spot catch trials.[6] There is one previous study by the inventors of the SITA family which suggested that SITA faster has a similar level of test–retest reproducibility compared with SITA fast.[6] Shortening of the VF measurement duration has obvious merits; it reduces the burden for patients, avoids inaccurate VF measurement due to fatigue, and enables clinical facilities to perform VF measurement in more patients at a given time.[7] However, test duration and accuracy are usually in a trade-off relation, and rapid measurement is often achieved by sacrificing accuracy. There is no VF measurement algorithm which is faster than SITA standard without losing test accuracy, to the best of our knowledge.

We previously constructed a model to VF progression prediction model using variational Bayes linear regression (VBLR).[8 9] VBLR is a method for predicting VF progression considering the spatial and temporal patterns of VF damage using the variational Bayes statistic. Prediction accuracy was far better compared with the conventional ordinary least squares linear regression model, such as mean deviation (MD) trend analysis. The possible usefulness of this model is not only for VF progression prediction, but also for VF measurement. VBLR can accurately predict VF sensitivities during VF measurement; each time VF sensitivity is decided, the prediction of VF sensitivities at remaining test points can be updated using VBLR. This accurate

prediction of VF sensitivity is expected to avoid showing redundant target presentation by optimising initial intensity in bracketing, and by interrupting the thresholding by applying a cut-off value of the likelihood of estimated VF sensitivity value. We constructed this VF test algorithm using VBLR (VBLR-VF) and validated its usefulness in patients with glaucoma. As a result, we achieved a shorter VF measurement time without losing test–retest reproducibility.

METHOD

Written consent was given by the patients for their information to be stored in the hospital database and used for research. This study was performed according to the tenets of the Declaration of Helsinki.

Subjects

This study included the 122 eyes of 73 patients with primary open-angle glaucoma (POAG), followed up at the Tokyo University Hospital. POAG was diagnosed when the following findings were present: (1) presence of typical glaucomatous changes in the optic nerve head (ONH), such as a rim notch with a rim width ≦0.1 disc diameters or vertical cup-to-disc ratio of >0.7 and/or a retinal nerve fibre layer defect with its edge at the ONH margin greater than a major retinal vessel, diverging in an arcuate or wedge shape, confirmed by a panel of glaucoma specialists (HM and RA) after inspecting photographs of the stereofundus; (2) presence of glaucomatous VF defects compatible with glaucomatous ONH changes, thus fulfilling Anderson-Patella's criterion[10] and (3) absence of other systemic or ocular disorders, including cataract, except for clinically insignificant senile cataract, shallow peripheral anterior chamber that could affect ONH and VF, and history of intraocular surgery or refractive surgery except for uneventful intraocular lens implantation. All patients were at least 20 years old.

VF measurement with SITA standard

The SITA standard measurement was carried out using the HFA 24–2 test and standard Goldmann III stimulus size.

VF measurement with VBLR-VF

VF measurement algorithm using VBLR was implemented in a perimetry by KOWA(AP-7700).

The VF measurement algorithm procedure is as follows: The prior model and starting stimulus intensities were calculated as the first-time measurement in each measurement in this investigation. In other words, this experiment was conducted as the initial VF measurement, and hence data from each patient's past VF measurement were not used.

Prior model

The Gaussian mixture distribution was set with VBLR using the dataset consisting of the 7268 eyes of 4278 subjects from the University of Tokyo Hospital.[8 9] The Gaussian mixture distribution was given by:

$$q(\mathbf{t}) = \sum_{k=0}^{K} \pi_k \mathcal{N}\left(\mathbf{t} \,|\, \mu_k, \Lambda_k^{-1}\right) \# (1)$$

where $\mathbf{t}$ is the total deviation values of the 52 test points with the 24–2 test grid, K is the number of the Gaussian distribution, π_k is the mixing coefficient in each Gaussian distribution, $\pi_k \in [0,1]$, $\sum_{k=0}^{K}\pi_k = 1$, μ_k is the average in each Gaussian distribution, and Λ_k is the precision matrix in each Gaussian distribution.

Starting stimulus intensities

In the case of a patient in whom VF measurement was carried out N times before, the VF measurement sequence was begun at $q(\mathbf{t}_{N+1})$ by equation 1 calculated with VBLR using the former N times results. In the case of a patient in whom VF measurement was carried out for the first time, the VF measurement sequence was begun at the level decided by the age-corrected normative database.

Strategy

Full Threshold strategy was used, that is, 4 dB until the first reversal, followed by 2 dB steps until the second reversal.

Model update through measurement

The VF measurement procedure was started by measuring the sensitivity at four primary points, one in each quadrant of the field at 12.7° from the point of fixation. After measuring the primary points, the Gaussian mixture distribution was updated when each point's sensitivity was decided. The Gaussian mixture distribution was divided into the points which the sensitivity was decided and the points which the sensitivity was not decided. The divided Gaussian mixture distribution was given by:

$$q(\mathbf{t_a} \,|\, \mathbf{t_b}) = \sum_{k=1}^{K} \hat{\pi}_k \mathcal{N}\left(\mathbf{t_a} \,|\, \mu_{a,k} - \Lambda_{aa,k}^{-1}\Lambda_{ab,\,k}\left(\mathbf{t_b} - \mu_{b,k}\right),\ \Lambda_{aa,k}^{-1}\right) \# (2)$$

$$\mathrm{t} = \begin{pmatrix} t_a \\ t_b \end{pmatrix}, \mu = \begin{pmatrix} \mu_a \\ \mu_b \end{pmatrix}, \Sigma = \begin{pmatrix} \Sigma_{aa} & \Sigma_{ab} \\ \Sigma_{ba} & \Sigma_{bb} \end{pmatrix}, \Lambda = \begin{pmatrix} \Lambda_{aa} & \Lambda_{ab} \\ \Lambda_{ba} & \Lambda_{bb} \end{pmatrix}$$

where $\mathbf{a}$ is the points which the sensitivity is not decided, $\mathbf{b}$ is the points the sensitivity was decided, $\hat{\pi}_k$ is updated the mixing coefficient in each Gaussian distribution, $\hat{\pi}_k \in [0,1]$, $\sum_{k=0}^{K}\hat{\pi}_k = 1$, Σ is the variance-covariance matrix in each Gaussian distribution.

Update starting stimulus intensities

The points which were not begun the VF measurement sequence were updated the starting stimulus intensities with $\mu_{a|b}$ by equation 2. By predicting appropriate sensitivities, the number of stimulus presentations could be reduced at most points.

Estimate sensitivities

When the sensitivity $\mu_{a|b}$ at the point was considered to have been estimated to sufficient accuracy with $\Sigma_{a|b}$ by equation 2, the VF measurement at the point was terminated. The level of accuracy was decided by the age-corrected normative database variance. The appropriate estimation could be reduced the number of stimulus presentations with a marginal decrease of measurement accuracy.

These VF measurements with HFA SITA standard and VBLR were conducted twice within 3 months period. The prior model and the starting stimulus intensities were calculated as the first-time measurement in each measurement in the current study.

Only reliable VFs were used in the analyses, defined as a fixation loss rate <33 %, a false-positive rate <33% and an FN rate <33%.

Statistical analysis

These values were compared between the VF measured with HFA SITA standard and VBLR-VF; (1) MD, (2) Pattern SD (PSD), (3) pointwise VF sensitivity and (4) measurement duration, using the linear mixed model where the random effect was patients. The linear mixed model is equivalent to ordinary linear regression in that the model describes the relationship between

the predictor variables and a single outcome variable. However, standard linear regression analysis makes the assumption that all observations are independent of each other. Measurements were nested within subjects and also test points in this study; hence, dependent of each other. Ignoring this grouping of the measurements will result in the underestimation of standard errors of regression coefficients. The linear mixed model adjusts for the hierarchical structure of the data, modelling in a way in which measurements are grouped within subjects to reduce the possible bias derived from the nested structure of data.[11 12]

With VBLR-VF, the MD and PSD were calculated as follow, similarly to HFA SITA standard.[10]

$$\mathrm{MD} = \left\{\frac{1}{n}\sum_{i=1}^{n}\frac{(X_i - N_i)}{S_i^2}\right\} \Big/ \left\{\frac{1}{n}\sum_{i=1}^{n}\frac{1}{S_i^2}\right\}$$

$$\mathrm{PSD} = \sqrt{\left\{\frac{1}{n}\sum_{i=1}^{n}S_i^2\right\} \cdot \left\{\frac{1}{n-1}\sum_{i=1}^{n}\frac{(X_i - N_i - MD)^2}{S_i^2}\right\}}$$

n: The number of the tested points.
X_i: The test result at the tested point (dB).
N_i : The age-corrected normal value at the tested point (dB).
S_i: The age-corrected SD at the tested point (dB).

In addition, the test–retest reproducibility of the two VF measurements were compared using 52 VF sensitivity values through the root mean squared error (RMSE) statistic, defined as follows:

$$\mathrm{RMSE} = \sqrt{\sum_{i=1}^{52}\frac{(\text{TD value of the } i\text{th point (1st VF)} - \text{TD value of the } i\text{th point (2nd VF)})^2}{52}}.$$

All analyses were performed using the statistical programming language 'R' (R V.3.1.3; The Foundation for Statistical Computing, Vienna, Austria).

RESULTS

Thirty-four of the 73 patients were male. Patients' age was 62.5±10.5 (39–84) (mean±SD, SD, (range)) years old. 62 in the 122 eyes were right eyes.

The MD value with the SITA standard was −7.9±6.2 (mean±SD, (range: −26.7 to 0.6]) in the first measurement and −8.7±6.2 (−27.3 to 0.3) dB in the second measurement, as shown in table 1. MD values with the VBLR-VF were −8.2±6.4 (−25.6 to 1.4) and −8.0±6.2 (−24.9 to 1.6) dB for the first and second measurements, respectively. There was no significant difference between these four MD values (p>0.05, linear mixed model with the adjustment for multiple comparisons using the Tukey's test). Figure 1 illustrates the relationship among MD values recorded with the SITA standard and the VBLR-VF. There were strong correlations in MD values between the first and second HFA SITA standard tests (correlation coefficient=0.97, p<0.0001, figure 1A) and also for the first and second VBLR tests (correlation coefficient=0.97, p<0.0001, figure 1B). In addition, there was a significant correlation between the MD values of the first HFA SITA standard tests and the first VBLR tests (correlation coefficient=0.97, p<0.0001, figure 1C) and

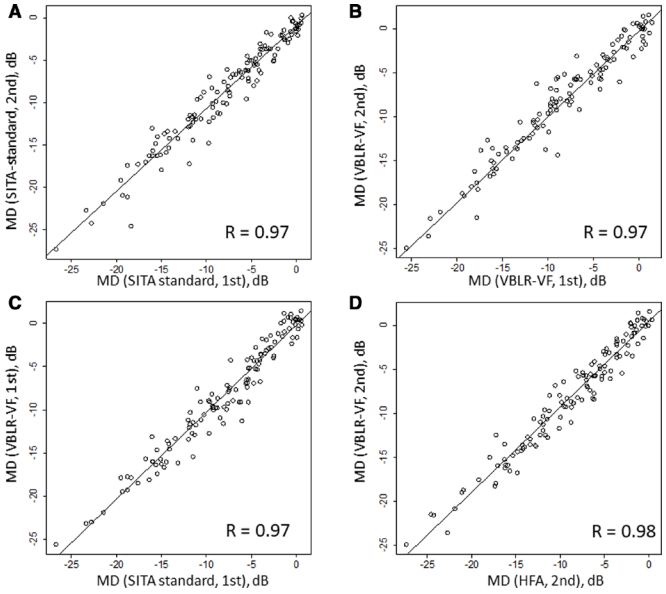

Figure 1 Relationship among MD values recorded with SITA standard and VBLR-VF. (A) MD values between first and second SITA standard, (B) MD values between first and second VBLR-VF, (C) MD values between first SITA standard and first VBLR-VF, (D) MD values between second SITA standard and second VBLR-VF. MD, mean deviation, SITA, Swedish interactive threshold algorithm, VBLR, variational Bayes linear regression, VF, visual field.

Table 1 Comparisons of the values of MD, PSD and pointwise sensitivities

variables		SITA-standard	VBLR-VF	P value*	P value**	P value†	P value††
1st measurement	MD, (mean±SD) (range), dB	−7.9±6.2 (−26.7 to 0.6)	−8.2±6.4 (−25.6 to 1.4)	0.99	0.99	0.96	0.99
2nd measurement	MD, (mean±SD) (range), dB	−8.7±6.2 (−27.3 to 0.3)	−8.0±6.2 (−24.9 to 1.6)				
1st measurement	PSD, (mean±SD) (range), dB	8.5±4.9 (1.1 to 16.0)	8.7±4.8 (0.9 to 16.1)	0.30	0.97	0.92	0.39
2nd measurement	PSD, (mean±SD) (range), dB	8.6±4.8 (1.1 to 16.6)	8.6±4.7 (0.9 to 16.2)				
1st measurement	Pointwise sensitivity, (mean±SD) (range), dB	21.8±11.2 (0 to 39)	21.0±11.4 (0 to 41)	1.00	0.64	<0.001	<0.001
2nd measurement	Pointwise sensitivity, (mean±SD) (range), dB	21.0±11.2 (0 to 35)	21.1±11.2 (0 to 40)				

*Shows the result of the comparison between SITA standard (1st) and SITA standard (2nd).
**Shows the result of the comparison between VBLR-VF (1st) and VBLR-VF (2nd).
†Shows the result of the comparison between SITA standard (1st) and VBLR-VF (1st).
††Shows the result of the comparison between SITA standard (2nd) and VBLR-VF (2nd).
MD, mean deviation; PSD, pattern SD; SITA, Swedish interactive threshold algorithm; VBLR, Variational Bayes Linear Regression; VF, visual field.

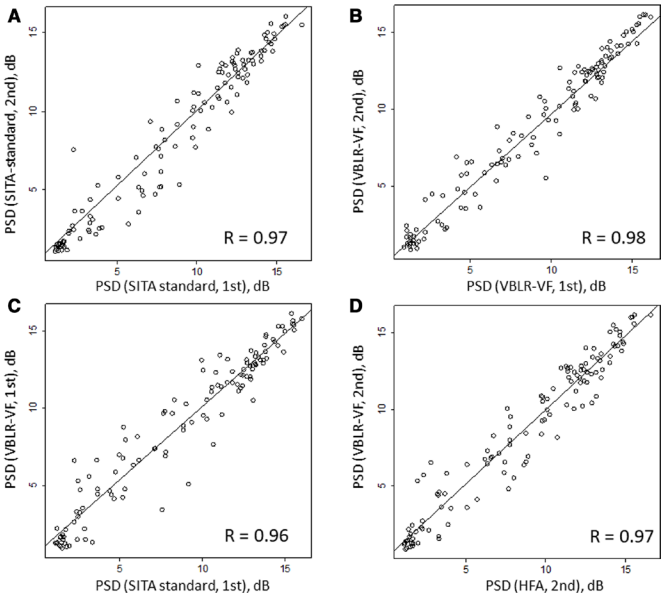

Figure 2 Relationship among PSD values recorded with the SITA standard and the VBLR-VF. (A) PSD values between first and second SITA standard, (B) PSD values between first and second VBLR-VF, (C) PSD values between first SITA standard and first VBLR-VF, (D) PSD values between second SITA standard and second VBLR-VF. HFA, Humphrey field analyser; PSD, pattern SD; SITA. Swedish interactive threshold algorithm; VBLR, variational Bayes linear regression; VF, visual field.

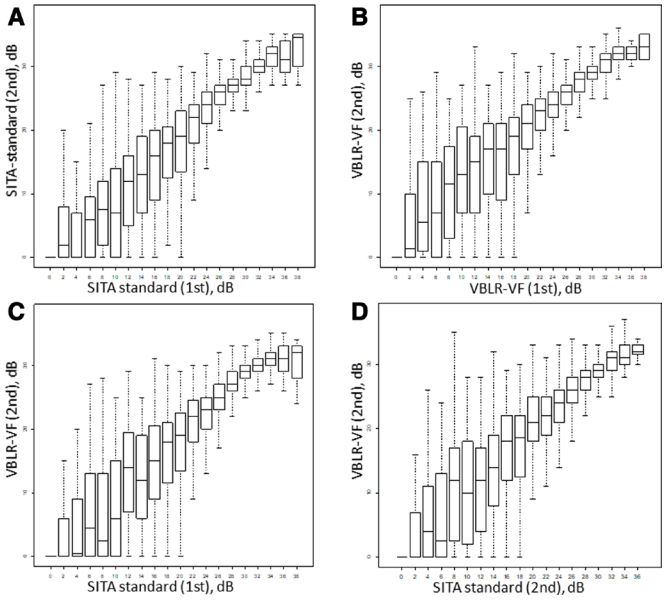

Figure 3 Relationship among pointwise VF sensitivities recorded with SITA standard and VBLR-VF. (A) Pointwise VF sensitivities between first and second SITA standard, (B) pointwise VF sensitivities between first and second VBLR-VF, (C) pointwise VF sensitivities between first SITA standard and first VBLR-VF, (D) pointwise VF sensitivities between second SITA standard and second VBLR-VF. SITA, Swedish interactive threshold algorithm; VBLR, variational Bayes linear regression; VF, visual field.

the MD values between their second tests (correlation coefficient=0.98, p<0.0001, figure 1D).

The PSD value with the HFA SITA standard test was 8.5±4.9 (mean±SD, (range: 1.1–16.0)) in the first measurement and 8.6±4.8 (1.1 to 16.6) dB in the second measurement, as shown in table 1. PSD values with the VBLR-VF were 8.7±4.8 (0.9 to 16.1) and −8.6±4.8 (0.9 to 16.2) dB for the first and second measurements, respectively. There was no significant difference between these four PSD values (p>0.05, linear mixed model). Figure 2 illustrates the relationship among PSD values recorded with the HFA and the VBLR. There were strong correlations in PSD values between the first and second HFA tests (correlation coefficient=0.97, p<0.0001, figure 2A) and also for the first and second VBLR-VF tests (correlation coefficient=0.98, p<0.0001, figure 2B). In addition, there was significant correlation between the PSD values of the first HFA SITA standard test and the first VBLR-VF test (correlation coefficient=0.96, p<0.0001, figure 2C) and the PSD values between their second tests (correlation coefficient=0.97, p<0.0001, figure 2D).

Table 1 shows the pointwise VF sensitivities in each measurement. The pointwise VF sensitivity values with the HFA SITA standard test was 21.8±11.2 (mean±SD, (range: 0–39)) in the first measurement and 21.0±11.2 (0–35) dB in the second measurement. Pointwise VF sensitivity values with the VBLR-VF were 21.0±11.4 (0–41) and 21.0±11.2 (0–40) dB for the first and second measurements, respectively. There was no significant difference in the pointwise VF sensitivity values between the first HFA SITA standard test and the second HFA SITA standard test, and also between the first VBLR-VF test and the second VBLR-VF test (p>0.05, linear mixed model with the adjustment for multiple comparisons using the Tukey's test), however, there was a significant difference in the pointwise VF sensitivity values between the first HFA SITA standard test and the first VBLR-VF test, and also between the second HFA SITA standard test and the second VBLR-VF test (p>0.05, linear mixed model with the adjustment for multiple comparisons using the Tukey's test). Figure 3A,B shows the test–retest agreements with SITA standard test and also VBLR-VF test in a fashion similar to that in Artes *et al*.[3] With both methods, precision decreased with increasing deficit. All of the pointwise VF sensitivities with SITA standard (first), SITA standard (second), VBLR-VF (first) and VBLR-VF (second) were significantly correlated to each other (all correlation coefficient=0.92, p<0.0001). These relationships are illustrated in figure 3C,D.

The RMSEs as a function of MD values of SITA standard test (first and second measurements were averaged) were illustrated

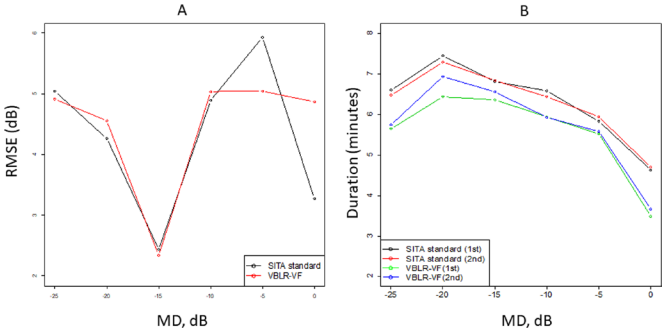

Figure 4 RMSE as a function of MD values. (A) The RMSE as a function of MD value is illustrated. MD value was calculated as the average of the first and second measurements with SITA-standard. (B) The measurement duration as a function of MD value is illustrated. MD value was calculated as the average of the first and second measurements with SITA-standard. MD, mean deviation; RMSE, root mean squared error; SITA, Swedish interactive threshold algorithm; VBLR, variational Bayes linear regression; VF, visual field.

in figure 4A. There was no significant difference in the RMSE values between SITA standard test and VBLR-VF test (p=0.68, linear mixed model). This was also the case when RMSEs were compared between the two methods in each of early (MD with SITA standard ≥6 dB), moderate (MD with SITA standard between −6 and −12 dB) and advanced stage (MD with SITA standard ≤12 dB)[13] (p=0.68, 0.66 and 0.53, respectively, linear mixed model).

The measurement duration with the HFA SITA standard test was 6 min and 2 s±1 min 6 s (3 min 57 s to 8 min 34 s) in the first measurement and 6 min 2 s±1 min 2 s (4 min to 8 min 27 s) in the second measurement. A significantly shorter measurement duration was observed for the VBLR-VF test: 5 min 23 s±1 min 32 s (2 min 49 s to 8 min 51 s) and 5 min 30 s±1 min 26 s (2 min 58 s to 8 min 44 s) for the first and second measurements, respectively (p<0.001). This was also the case when the measurement durations were compared between the two methods in each of early (MD with SITA standard ≥6 dB), moderate (MD with SITA standard between −6 and −12 dB) and advanced stage (MD with SITA standard ≤12 dB)[13] (all p<0.001 except for eyes in advanced stage in the second measurement: p=0.013, linear mixed model). The measurement duration was illustrated in the relationship to the MD values of SITA standard test (first and second measurements were averaged), figure 4B.

DISCUSSION

VF measurement was conducted in this study, using a novel algorithm of VBLR-VF, and the measurement results were compared with the SITA standard programme with HFA, in eyes with POAG. As a result, the measured sensitivity values were very closely related between the VBLR-VF and the SITA standard methods. More specifically, MD and PSD values had the correlation coefficient values between 0.97 and 0.99 (figures 1 and 2), and there were no significant differences in the MD and PSD values between SITA standard and VBLR-VF (table 1). The pointwise sensitivities were also very closely related between SITA standard and VBLR-VF with the correlation coefficient of 0.92. There was a significant difference in the pointwise sensitivities between SITA standard and VBLR-VF, however, there was no systematic trend which algorithm is associated with higher sensitivity than another in the repeated measurements (VF sensitivity was higher and lower with the SITA standard in the first and second measurement, respectively). The RMSE value of the test–retest reproducibility was not significantly different between the two methods. In contrast, the measurement duration was significantly shorter with the VBLR method by 30 or 40 s. The reduction of the VF measurement duration was observed across the disease status, as illustrated in figure 4B.

Since the algorithm of SITA was introduced in HFA in 1980s, despite efforts, there has been no method which succeeded to shorten the testing duration without losing the accuracy. The currently proposed VBLR-VF algorithm achieved significantly faster VF measurement in the initial VF measurement, compared with SITA standard. Of note, it was not associated with the increase of the test–retest reproducibility, unlike other SITA families, such as fast and faster.[3 5 6] This is probably because SITA fast and faster attempted a faster measurement simply by omitting some reversals in the thresholding and also false negative test, whereas VBLR-VF avoided redundant target presentation by accurately predicting threshold at each test point using VBLR. In the bracketing thresholding method, the time required to determine the threshold becomes long when the luminance of the initial target presentation is far different from the threshold, because a larger number of target presentations are needed. In SITA standard, thresholds of the adjacent test points are predicted using the maximum likelihood estimation derived from a database of glaucomatous and normative VFs, so that the initial luminance level is optimisation. In addition, in SITA standard, thresholding is even terminated earlier using the Bayesian posterior probability of the threshold. The currently proposed VBLR-VF algorithm takes similar approach with SITA standard that the optimisation of the initial luminance level and early termination are conducted, so that VF measurement duration becomes short. In addition, both SITA standard and VBLR predict VF sensitivities using the Bayes statistic. In VBLR, the prediction of VF sensitivity was given, considering the spatial and temporal patterns of the VF damage and also the correlation across the test points.[8 9] As a result, we have reported that very accurate prediction of future VF progression can be made; the prediction accuracy when predicting 10th future VF even with the initial 1 or 2 VFs using the VBLR was almost identical to that with 9 VFs using ordinary linear regression. In the current study, this accurate prediction also enabled significantly faster measurement of VF compared with SITA standard.

Attempts have been made to shorten the VF measurement also in perimetries other than HFA. The German Adaptive Thresholding Estimation (GATE) in Octopus perimetry (Haag-Streit, Switzerland) is one of such examples, in which the initial target luminance is decided considering the patient's past VF.[14 15] A previous study suggested that this approach resulted in identical or slightly poorer (where VF sensitivity is high) test–retest reproducibility with GATE compared with SITA standard, with a shorter measurement time (approximately 16% reduction: 4.7 min with GATE and 5.6 min with SITA standard in average). This shortening effect was slightly more than the currently proposed VBLR-VF algorithm shortened approximately 11 or 8% compared with SITA standard. However, a merit of the VBLR-VF algorithm over the GATE programme is that previous VF measurement is not needed, and can be applied to even the initial VF measurement. In addition, the advantage of GATE measurement may be questionable when very long period, such as 5 or 10 years, has passed from the prior VF measurement, because of the possible large discrepancy between the past and current VFs. The current experiment was conducted without using patients' past VF data, assuming an initial VF measurement in each patient. In other words, more accurate prediction would be given, when the patient's past VF data is also utilised in the VF sensitivity prediction using VBLR. It would be of interest to investigate in future whether VBLR-VF can shorten VF measurement no less than GATE programme under this condition. There are other programmes in Octopus perimetry, such as Dynamic strategy and Tendency Oriented Perimetry programme. These were developed aiming at fast VF measurement, by varying the luminance interval in the bracketing according to the eccentricity from fovea (Dynamic strategy) or performing target presentation only once at each test point and inferring adjacent test points' thresholds.[16] These measurements are known to be shorter than SITA standard measurement; however, the measurement accuracies have been reported to be optimal.[17–20] Inaccurate measurement of VF sensitivity is not only problematic in assessing the disease status of glaucoma, but also in detecting progression, because detection of progression is delayed associated with larger variability of VF measurement, as reported by Jansonius.[21] The zippy estimation by sequential testing (ZEST)[22–24] is another VF measurement algorithm which has been reported as possibly useful

to measurement accuracy compared with SITA, but not the measurement duration.[25] A further merit of ZEST is that it is available for use on an Octopus perimeter through the Open Perimetry Initiative, where the source code is also available, unlike VBLR-VF. A study would be needed to compare the measurement accuracy and duration between the currently proposed VBLR-VF and ZEST in future.

Jansonius reported that the detection of progression does not only rely on the accuracy (variability), but also the frequency of VF measurement; progression detection is delayed when less frequently measured.[21] SITA-standard algorithm enabled much faster VF measurement compared with its Full Threshold algorithm predecessor, however previous studies have reported that it is difficult enough for busy clinics to carry out central VF testing with sufficient frequency.[26 27] Recent studies revealed the importance of VF measurement in the central 10 degrees,[28–32] such as HFA 10–2 test, has further postulated this aspect. Faster VF measurement without losing the measurement accuracy, such as with the VBLR-VF algorithm, would be needed to resolve the problem of the burden of the VF measurement to clinical facilities as well. Thus the possible merit of the faster VF measurement with fast VF measurement is not only in the reduction of the patients' burden and prevention of fatigue effect. As shown in the current study, the measurement duration with VBLR-VF was significantly faster than that of SITA standard, however, the difference was merely less than 1 min which is relatively small compared with the shortening effect with SITA fast or faster, and may not be sufficient to prevent patients' fatigue considerably. One of the possible reasons would be that the current study was conducted assuming the first VF measurement for each eye. A further reduction of the measurement can be expected when VF measurement is conducted thereafter, because VF prediction in VBLR can be much more accurate when the previous VF record of the patient is used. A further study should be conducted shedding light on this issue.

There are several limitations to this study. As VBLR can predict patients' VFs accurately when past VF data are used, nonetheless this information was not used in this study. This was because the current study was conducted assuming the initial VF measurement in each patient, however a further faster VF measurement may be achieved through more accurate prediction in VBLR with the patient's previous VF data, which should be investigated in a future study. In addition, current study was conducted using the HFA 24–2 test grid. The usefulness of this approach should be evaluated with the HFA 10–2 test grid.

In conclusion, we developed a novel VF measurement algorithm of VBLR-VF. This programme shortened VF measurement by 30 or 40 s, compared with the SITA standard programme without losing measurement accuracy.

Contributors RA, HM and SS researched literature and conceived the study. RA, HM and NS were involved in protocol development and gaining ethical approval. All authors were involved in patient recruitment and data analysis. RA wrote the first draft of the manuscript. All authors reviewed and edited the manuscript and approved the final version of the manuscript.

Funding Supported in part by Japan Science and Technology Agency (JST) CREST (RA, HM), Grants 00768351 (MM), 20 768 254 (YF), 25 861 618 (HM) and 19H01114, 18KK0253, and 26 462 679 (RA) from the Ministry of Education, Culture, Sports, Science and Technology of Japan, grant from the University of Tokyo Gap Fund ProgramProgramme (no grant number), and grants from Suzuken Memorial Foundation (no grant number) and Taiju Life Social Welfare Foundation (Mitsui Life Social Welfare Foundation) (no grant number).

Competing interests None declared.

Patient consent for publication Not required.

Ethics approval This study was approved by the Research Ethics Committee of the Graduate School of Medicine and Faculty of Medicine at the University of Tokyo.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available on reasonable request.

ORCID iDs
Ryo Asaoka http://orcid.org/0000-0001-7182-1912
Yuri Fujino http://orcid.org/0000-0001-6082-0738
Masato Matsuura http://orcid.org/0000-0002-6120-1100
Kazunori Hirasawa http://orcid.org/0000-0001-5100-005X

REFERENCES

1 Bengtsson B, Olsson J, Heijl A, *et al*. A new generation of algorithms for computerized threshold perimetry, SITA. *Acta Ophthalmol Scand* 1997;75:368–75. doi:10.1111/j.1600-0420.1997.tb00392.x
2 Bengtsson B, Heijl A. Evaluation of a new perimetric threshold strategy, SITA, in patients with manifest and suspect glaucoma. *Acta Ophthalmol Scand* 1998;76:268–72. doi:10.1034/j.1600-0420.1998.760303.x
3 Artes PH, Iwase A, Ohno Y, *et al*. Properties of perimetric threshold estimates from full threshold, SITA standard, and SITA fast strategies. *Invest Ophthalmol Vis Sci* 2002;43:2654–9.
4 Bengtsson B, Heijl A, Fast S. SITA fast, a new rapid perimetric threshold test. description of methods and evaluation in patients with manifest and suspect glaucoma. *Acta Ophthalmol Scand* 1998;76:431–7. doi:10.1034/j.1600-0420.1998.760408.x
5 Saunders LJ, Russell RA, Crabb DP. Measurement precision in a series of visual fields acquired by the standard and fast versions of the Swedish interactive thresholding algorithm: analysis of large-scale data from clinics. *JAMA Ophthalmol* 2015;133:74–80. doi:10.1001/jamaophthalmol.2014.4237
6 Heijl A, Patella VM, Chong LX, *et al*. A new SITA perimetric threshold testing algorithm: construction and a multicenter clinical study. *Am J Ophthalmol* 2019;198:154–65. doi:10.1016/j.ajo.2018.10.010
7 Johnson CA, Adams CW, Lewis RA. Fatigue effects in automated perimetry. *Appl Opt* 1988;27:1030–7. doi:10.1364/AO.27.001030
8 Murata H, Araie M, Asaoka R. A new approach to measure visual field progression in glaucoma patients using variational Bayes linear regression. *Invest Ophthalmol Vis Sci* 2014;55:8386–92. doi:10.1167/iovs.14-14625
9 Murata H, Zangwill LM, Fujino Y, *et al*. Validating variational Bayes linear regression method with Multi-Central datasets. *Invest Ophthalmol Vis Sci* 2018;59:1897–904. doi:10.1167/iovs.17-22907
10 Anderson D, Patella V. *Automated static perimetry*. 2nd edn. St. Louis: Mosby, 1999.
11 Baayen RH, Davidson DJ, Bates DM. Mixed-effects modeling with crossed random effects for subjects and items. *J Mem Lang* 2008;59:390–412. doi:10.1016/j.jml.2007.12.005
12 Bates D, Mächler M, Bolker B, *et al*. Fitting Linear Mixed-Effects Models Using lme4. *J Stat Softw* 2015;67. doi:10.18637/jss.v067.i01
13 Anderson DR, Patella VM. *Automated static perimetry*. 2nd edn. St.Louis: Mosby, 1999.
14 Luithardt AF, Meisner C, Monhart M, *et al*. Validation of a new static perimetric thresholding strategy (gate). *Br J Ophthalmol* 2015;99:11–15. doi:10.1136/bjophthalmol-2013-304535
15 Schiefer U, Pascual JP, Edmunds B, *et al*. Comparison of the new perimetric gate strategy with conventional full-threshold and SITA standard strategies. *Invest Ophthalmol Vis Sci* 2009;50:488–94. doi:10.1167/iovs.08-2229
16 Gonzalez de la Rosa M, Gonzalez-Hernandez M, Garcia Feijoo J, *et al*. Diagnostic accuracy and reproducibility of tendency oriented perimetry in glaucoma. *Eur J Ophthalmol* 2006;16:259–67–67. doi:10.1177/112067210601600211
17 Anderson AJ. Spatial resolution of the tendency-oriented perimetry algorithm. *Invest Ophthalmol Vis Sci* 2003;44:1962–8. doi:10.1167/iovs.02-0828
18 King AJW, Taguri A, Wadood AC, *et al*. Comparison of two fast strategies, SITA fast and top, for the assessment of visual fields in glaucoma patients. *Graefes Arch Clin Exp Ophthalmol* 2002;240:481–7. doi:10.1007/s00417-002-0482-y
19 Maeda H, Nakaura M, Negi A. New perimetric threshold test algorithm with dynamic strategy and tendency oriented perimetry (top) in glaucomatous eyes. *Eye* 2000;14:747–51. doi:10.1038/eye.2000.196
20 Morales J, Weitzman ML, González de la Rosa M. Comparison between Tendency-Oriented perimetry (top) and octopus threshold perimetry. *Ophthalmology* 2000;107:134–42. doi:10.1016/S0161-6420(99)00026-3
21 Jansonius NM. On the accuracy of measuring rates of visual field change in glaucoma. *Br J Ophthalmol* 2010;94:1404–5. doi:10.1136/bjo.2009.164897

22 King-Smith PE, Grigsby SS, Vingrys AJ, *et al*. Efficient and unbiased modifications of the quest threshold method: theory, simulations, experimental evaluation and practical implementation. *Vision Res* 1994;34:885–912. doi:10.1016/0042-6989(94)90039-6

23 Vingrys AJ, Pianta MJ. A new look at threshold estimation algorithms for automated static perimetry. *Optom Vis Sci* 1999;76:588–95. doi:10.1097/00006324-199908000-00028

24 Turpin A, McKendrick AM, Johnson CA, *et al*. Performance of efficient test procedures for frequency-doubling technology perimetry in normal and glaucomatous eyes. *Invest Ophthalmol Vis Sci* 2002;43:709–15.

25 Turpin A, McKendrick AM, Johnson CA, *et al*. Properties of perimetric threshold estimates from full threshold, ZEST, and SITA-like strategies, as determined by computer simulation. *Invest Ophthalmol Vis Sci* 2003;44:4787–95. doi:10.1167/iovs.03-0023

26 Malik R, Baker H, Russell RA, *et al*. A survey of attitudes of glaucoma subspecialists in England and Wales to visual field test intervals in relation to NICE guidelines. *BMJ Open* 2013;3. doi:10.1136/bmjopen-2012-002067. [Epub ahead of print: 03 May 2013].

27 Crabb DP, Russell RA, Malik R, *et al*. Frequency of visual field testing when monitoring patients newly diagnosed with glaucoma: mixed methods and modelling. *Health Serv Deliv Res* 2014;2:1–102. doi:10.3310/hsdr02270

28 De Moraes CG, Hood DC, Thenappan A, *et al*. 24-2 visual fields miss central defects shown on 10-2 tests in glaucoma suspects, ocular hypertensives, and early glaucoma. *Ophthalmology* 2017;124:1449–56. doi:10.1016/j.ophtha.2017.04.021

29 Hood DC, Raza AS, de Moraes CGV, *et al*. Glaucomatous damage of the macula. *Prog Retin Eye Res* 2013;32:1–21. doi:10.1016/j.preteyeres.2012.08.003

30 Park H-YL, Hwang B-E, Shin H-Y, *et al*. Clinical clues to predict the presence of parafoveal scotoma on Humphrey 10-2 visual field using a Humphrey 24-2 visual field. *Am J Ophthalmol* 2016;161:150–9. doi:10.1016/j.ajo.2015.10.007

31 Tomairek RH, Aboud SA, Hassan M, *et al*. Studying the role of 10-2 visual field test in different stages of glaucoma. *Eur J Ophthalmol* 2020;30:706–13. doi:10.1177/1120672119836904

32 Traynis I, De Moraes CG, Raza AS, *et al*. Prevalence and nature of early glaucomatous defects in the central 10° of the visual field. *JAMA Ophthalmol* 2014;132:291–7. doi:10.1001/jamaophthalmol.2013.7656