Author/s:
Muthusamy, V;Turpin, A;Walland, MJ;Nguyen, BN;McKendrick, AM

Title:
Increasing the Spatial Resolution of Visual Field Tests Without Increasing Test Duration: An Evaluation of ARREST

Date:
2020-12

Citation:
Muthusamy, V., Turpin, A., Walland, M. J., Nguyen, B. N. & McKendrick, A. M. (2020). Increasing the Spatial Resolution of Visual Field Tests Without Increasing Test Duration: An Evaluation of ARREST. Translational Vision Science and Technology, 9 (13), pp.1-13. https://doi.org/10.1167/tvst.9.13.24.

Persistent Link:
https://hdl.handle.net/11343/262688

tvst

Article

Increasing the Spatial Resolution of Visual Field Tests Without Increasing Test Duration: An Evaluation of ARREST

Vasanth Muthusamy[1], Andrew Turpin[2], Mark J. Walland[3], Bao N. Nguyen[1], and Allison M. McKendrick[1]

[1] Department of Optometry and Vision Sciences, The University of Melbourne, Parkville, Victoria, Australia
[2] School of Computing and Information Systems, The University of Melbourne, Parkville, Victoria, Australia
[3] Department of Medical Education (St. Vincent's Hospital Clinical School), The University of Melbourne, Parkville, Victoria, Australia

Correspondence: Allison M. McKendrick, Department of Optometry and Vision Sciences, The University of Melbourne, Grattan Street, Parkville, Victoria 3010, Australia. e-mail: allisonm@unimelb.edu.au

Received: August 11, 2020
Accepted: October 25, 2020
Published: December 16, 2020

Keywords: perimetry; visual fields; test grids; spatial resolution; test duration

Citation: Muthusamy V, Turpin A, Walland MJ, Nguyen BN, McKendrick AM. Increasing the spatial resolution of visual field tests without increasing test duration: An evaluation of ARREST. Trans Vis Sci Tech. 2020;9(13):24, https://doi.org/10.1167/tvst.9.13.24

Purpose: The Australian Reduced Range Extended Spatial Test (ARREST) approach was designed to improve visual field spatial resolution while maintaining a similar test duration to clinically used testing algorithms. ARREST does not completely threshold visual field locations with sensitivity < 17 dB, and uses the presentations saved to test new locations in areas of steep gradient within the visual field. Previous assessments of ARREST's performance have used computer simulation. In this study, we cross-sectionally assessed the performance of ARREST in people with visual field loss.

Methods: We tested 23 people with glaucoma (mean age: 71 $\pm$ 8 years) with established visual field loss. Three visual field procedures were performed using the Open Perimetry Interface: cZEST and ARREST on the Octopus 900 perimeter (Haag-Streit AG, Switzerland), and a reference standard (best available estimate [BAE]) on the Compass perimeter (CenterVue SpA, Italy). ARREST was compared against the cZEST and the BAE.

Results: On average, ARREST added seven new locations (range = 0–15) to a visual field test. There was no significant difference in the number of stimulus presentations between procedures (mean = 259 $\pm$ 25 [ARREST] vs. 261 $\pm$ 25 [cZEST], $P = 0.78$). In classifying threshold values < 17 dB, ARREST performed similarly when compared against BAE.

Conclusions: This study provides empirical evidence to support conclusions from previous computer simulations that ARREST can be used to increase spatial sampling in regions of interest without increasing test time.

Translational Relevance: ARREST is a new approach that augments current visual field testing procedures to provide better spatial description of visual field defects without increasing test duration.

Introduction

Static Automated Perimetry (SAP) typically estimates visual field sensitivity on a fixed test grid of locations regardless of the disease severity of the patient. Depending on the commercially available perimeter, test grids are either arranged in a cartesian or polar grid. For example, the 24-2 test pattern in the Humphrey Field Analyzer (HFA; Carl Zeiss Meditec, Dublin, CA) is cartesian, and the G-pattern in the Octopus 900 (Haag-Streit AG, Koeniz, Switzerland) is polar. Test locations are approximately 6 degrees apart in the visual field, and the full test typically requires between 5 to 10 minutes.[1,2] Recently, the Australian Reduced Range Extended Spatial Test (ARREST)[3] approach was designed to improve the spatial resolution of fixed test grids while preserving all other current attributes of SAP, such as test duration and ability to detect change in the visual field.

ARREST is designed to improve upon three limitations of current visual field testing procedures: (1) to increase spatial sampling in localized areas of interest in the visual field; (2) to avoid unnecessary testing of

locations that are known to be perimetrically "blind"; and (3) to minimize repeated stimulus presentations in areas of known high variability.

Regarding the first limitation, the sparsely arranged standard 6 degree spaced test grid is typically used clinically despite the nature of spatial visual field loss varying markedly between people.[4–6] Studies have shown that when locally condensed test grids are applied on morphologically suspicious glaucomatous areas, a greater number of defects are identified[7] and the detection rate of visual field progression is improved compared to a fixed test grid.[8,9] However, increasing the number of test locations also increases test duration, which may not be suitable for clinical testing.[8] ARREST is designed to add additional locations in areas of steep visual field gradient without increasing test duration.[3]

Second, current test procedures repeatedly assess perimetrically blind locations at each subsequent test visit. Testing locations that are well established to be at the measurement floor (0 dB) provides no additional information.[10] Retesting these locations during follow-up visits is time-consuming, and in people with moderate to advanced visual field loss, long durations of unseen stimuli may increase anxiety and frustration with visual field assessment. ARREST does not retest locations with established sensitivity of < 0 dB.[3]

The third limitation of current procedures is the well-studied higher test-retest variability in moderate and advanced visual field loss,[11–16] which reduces the accuracy of sensitivity estimates. Recently, Gardiner et al.[16] have shown that testing locations with sensitivity below approximately 15 to 19 dB with current SAP implementations does not provide information that is particularly clinically useful due to the degree of measurement noise. For example, it has been demonstrated that if sensitivity values below 19 dB are censored, the ability to classify visual field series as progressing using point-wise linear regression analysis is unchanged.[17] Nevertheless, current test procedures expend a lot of test presentations providing such noisy sensitivity estimates in areas of visual field damage.

ARREST takes a different approach. Rather than fully thresholding these locations, once a location has been established during a prior visual field test to have sensitivity < 17 dB, future tests simply check whether the location is still not perimetrically blind (able to see a 0 dB stimulus). Locations with sensitivity < 17 dB undergo a series of confirmatory stimulus presentation checks to classify threshold values at each test location as confirmed blind "< 0 dB (red)", likely blind "0 dB (orange)", and highly variable "1 dB to 16 dB (yellow)." The saved presentations are used to increase spatial density to assess new locations adjacent to established scotomas.[3] For locations where sensitivity is greater than 17 dB, normal thresholding approaches apply. In summary, the ARREST approach only modifies the current test grid and test algorithm when there are visual field sensitivities < 17 dB in the previous visual field test.

Previously, we have demonstrated the potential of the ARREST approach via computer simulation,[3] because the advantages of the approach are best visualized over longitudinal series of visual field tests. In simulation, the performance of ARREST was comparable to Zippy Estimation by Sequential Testing (ZEST) on a 24-2 test grid in detecting visual field progression, while improving spatial resolution with 25% to 40% fewer stimulus presentations than the ZEST 24-2 test procedure.[3]

The current study was designed as a feasibility study of the ARREST approach in people with glaucomatous visual field loss, to determine whether the main predictions from computer simulation are realized in human testing. We compared the performance of ARREST against ZEST (a variant of procedures used in commercial perimeters). The main advantages of the ARREST approach are likely to be realized within longitudinal series, hence to simulate the beginning of a test series we selected eyes with known visual field loss less than 17 dB and tested twice. We also include data from several case examples of people tested one year after their baseline visit. We tested people with glaucoma because it is a well-established condition that results in visual field loss but the ARREST approach is not specific to glaucomatous visual field loss. The main advantages of ARREST are visualization of increased spatial sampling of visual field deficits, however, we also explore the feasibility of calculating a global index like mean deviation[18] for the ARREST approach, which may be of some use as a clinical descriptor.

Methods

Study Participants

We tested 23 participants with glaucoma (mean age $\pm$ SD: 71 $\pm$ 8 years). Inclusion criteria included: an ophthalmological diagnosis of glaucoma, visual field sensitivities less than 17 dB in at least 3 test locations on their most recent HFA 24-2 SITA standard examination in at least one eye and retinal nerve fiber layer thinning outside normal limits measured using optical coherence tomography (OCT). Participants were referred from author M.W. or were previous participants in perimetry related studies in our department. Participants were excluded if the best-corrected

visual acuity was worse than 6/9.6 in the test eye, spherical equivalent refractive error greater than +6 or less than –6 diopters, any anterior segment media opacities and significant cataract on clinical examination, history of neurological disorders, or retinal pathology other than glaucoma that could influence the outcome of visual field assessment. The mean total deviation (MTD) of the cohort is discussed below.

Eighteen visually normal people, of similar age to the glaucoma group (mean ± SD: 68 ± 6 years) were recruited to create a normal reference range of performance for each location to compute MTD for the specific visual field algorithms used in this study. The same inclusion and exclusion criteria were followed with the addition of excluding people with a history of glaucoma or suspicious of developing glaucoma. All participants voluntarily provided written informed consent to participate in the study. This study adhered to the tenets of the Declaration of Helsinki and was approved by the research ethics committee of the University of Melbourne Human Research Ethics Committee (Ethics ID: 164695.7).

Overview of Visual Field Testing

All visual field algorithms were implemented using the Open Perimetry Interface (OPI).[19] ARREST was engineered as in the simulation study,[3] and a baseline ZEST procedure on a 24-2 grid was engineered to represent current SAP algorithms used in clinical settings (details below); we refer to it as cZEST, for "clinical ZEST." Both the ARREST and the cZEST tests were performed on the Octopus 900 perimeter (Haag-Streit AG, Koeniz, Switzerland). We also measured an additional best available estimate (BAE) to be used as a "ground truth" to compare against the results of both ARREST and cZEST. The BAE was an average of two visual field tests on a two degree separated test grid using a theoretically more accurate ZEST[20] procedure than cZEST (further details of each algorithm are below). The BAE procedure was also implemented using the OPI, but on the Compass perimeter (CenterVue, Padova, Italy), which uses fundus tracking.

For all visual field tests, we used Goldmann Size III stimuli on a fixed background intensity of 31.4 apostilbs (10 cd/m2). A luminance of 10,000 apostilbs was used as 0 dB. False-positive rates were measured using catch trials at every 30 stimulus presentations.[21] Only one eye was tested for each participant and the laterality was decided based on the above-mentioned visual field criteria. If both eyes satisfied the inclusion criteria, the eye with a higher number of damaged locations < 17 dB on the most recent HFA SITA standard 24-2 field was chosen. Left eye data were converted to right eye format for data analysis. The testing sequence of ARREST and cZEST was randomized and followed by BAE. Adequate breaks were given between the procedures.

Visual Field Algorithms

cZEST

Generally, the ZEST procedure refines a prior distribution of population threshold values (probability density function [pdf]) to estimate a final threshold and can be parametrized in many different ways.[20] The details of different ZEST procedures used in the context of perimetry is described in full in some previous papers.[22,23] Usually commercial perimeters seed a starting stimulus for test locations based on their neighbor's thresholds with the intent of saving test duration.[24] For our clinical baseline cZEST, therefore, we implement the "growth pattern" technique used by the HFA[24] on a 24-2 pattern to mimic current clinical scenario (Fig. 1 in Turpin et al.).[23]

In this study, the test starts by assigning a fixed bi-modal prior with the mean of the normal peak centered at 25 dB for the 4 primary test locations (one location in each quadrant positioned at $x = \pm 9$ and $y = \pm 9$ in cartesian coordinates). Once the threshold is estimated at these four locations, the pdf is modified in the neighboring locations based on the estimated threshold by setting the mode of the prior to an average of neighboring value(s). The thresholding process is repeated until the SD of the pdf is < 1.5 dB. The final threshold sensitivity estimated is the mean of the final pdf for that location. The maximum number of stimulus presentations at each location was capped at 12.

ARREST

The ARREST approach can be implemented with any currently available clinical perimetric algorithm and is outlined in Figure 1. In this study, we used the same procedure as in the previous simulation work where the underlying test was ZEST on the standard 24-2 test grid.[3] Note that the first ARREST test does not add any new locations to the 24-2 grid, and so we needed to do both an initial test (ARREST 1) and a follow-up test (ARREST 2) on the same day so that we could evaluate ARREST in our cross-sectional experiment. In a real scenario, these visits would be conducted some time apart. That is, the underlying test procedure (in this case, ZEST but could be any extant visual field thresholding algorithm) would be performed as usual at visits until some test locations fall below 17 dB, which would then trigger the ARREST

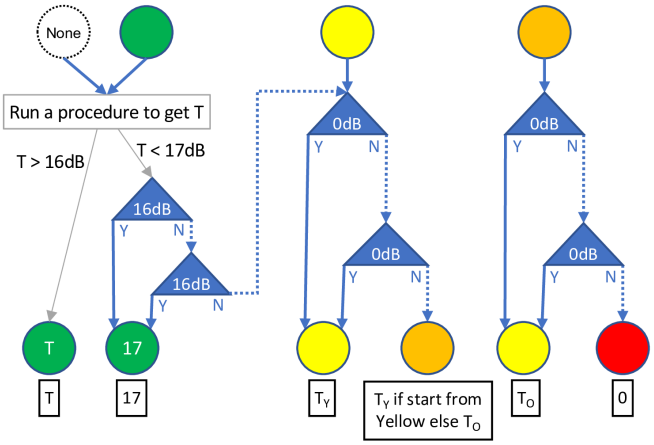

Figure 1. Steps in ARREST for thresholding one location. The *top row* represents the location's result from a previous test, and the *bottom circles* the result from this test. *Blue triangles* indicate presentations, with a *dotted line* being the path for an "unseen" (no) response, and *solid line* a path for a "seen" (yes) response. The numbers in *boxes* underneath represent the values used for computing mean total deviation, MTD (see text for details), but T_Y and T_O are not displayed to clinicians. "T" stands for test result of the underlying procedure (ZEST in this study). A red location remains red and receives no presentations.

decision rules to be additionally applied to subsequent tests. The underlying ZEST algorithm is as for cZEST without a limit of 12 presentations per location and without a growth pattern (as used in our prior simulation study of ARREST performance) [3]; thus, the mode of the prior pdf was at 25 dB as for all locations. These differences are enough to change normative values in some locations, hence we collected separate normative data for cZEST and ARREST.

ARREST sets an upper bound on the required number of stimulus presentations (250 in this study), and estimates the number and exact position of additional test locations for the current test based on the previously available visual field report. ARREST estimates the required number of stimulus presentation for the subsequent visit by assuming 10, 2, 1, and 0 presentations for green, yellow, orange, and red locations, respectively. If the sum of the assumed number of presentations is < 250, ARREST adds new locations assuming 10 presentations for each new location until the budget of 250 presentations is expended. The new locations are placed between a yellow location and a green location that has the highest threshold compared to its neighbors.

The output (Fig. 2, left side) from ARREST consists of locations that have a measured sensitivity ("green"), or are flagged as: "yellow", if the estimated sensitivity is below 17 dB and 0 dB was seen; "orange", if the estimated threshold is 0 dB for the first time on the day of testing; or "red" if the location was previously red, or previously orange, and 0 dB was not seen twice in this test. Red locations are excluded from testing in future visits. If any test location is flagged as "yellow" (less than 17 dB) it will stay as "yellow" and may progress to "orange" or "red" if 0 dB is not seen in subsequent tests.

The Best Available Estimate

As ARREST may test locations not on the 24-2 grid, in order to assess the accuracy of both cZEST and ARREST, we collected estimates of thresholds at a higher spatial resolution. Testing an entire central visual field with a high-resolution test grid is very time-consuming, and thus very difficult, particularly with elderly participants. Instead, for each participant, we chose one of five separate high-resolution test grids (supero-temporal, supero-nasal, infero-temporal, infero-nasal, and macula) having 100 test locations each spaced 2 degrees apart (Fig. 2, right) extending to 20 degrees on either side of the field in each quadrant except for the macular test grid, which covered ± 10 degrees from fixation. Two participants did not perform the visual field procedure using the compass (1 did not participate for the compass session and 1 had small pupils and postural difficulty) resulting in a total of 21 people for BAE analysis.

For each participant, we selected the high-resolution test grid that covered the area in the ARREST field that had the maximum number of newly added locations. If the ARREST field produced no additional locations (one participant), the macular grid was chosen. Each sensitivity in the chosen grid was measured using a ZEST with a uniform prior over the domain -5 to 40 dB, the same likelihood functions as in cZEST, and stopping after the SD of the pdf was < 1.5 dB.[25] The uniform prior assumes equal probability for each possible thresholds and thus avoids bias caused by the bi-modal[23] prior used in the ARREST and cZEST implementations.

For determining the BAE, if we were to test 100 locations in a single session with the uniform prior, the test duration would be approximately 20 to 25 minutes, depending upon the defect depth and the total area of the glaucomatous defect. To reduce the potential for errors caused by fatigue,[26] we split the 100 test locations into 4 sub-grids (25 locations evenly spaced) and each sub-grid was tested in a random order. Between each sub-grid, an adequate rest break was given. The procedure was repeated twice in the same day or on a separate day within 3 months depending on the stamina and preference of the participant. The OPI interface for the compass

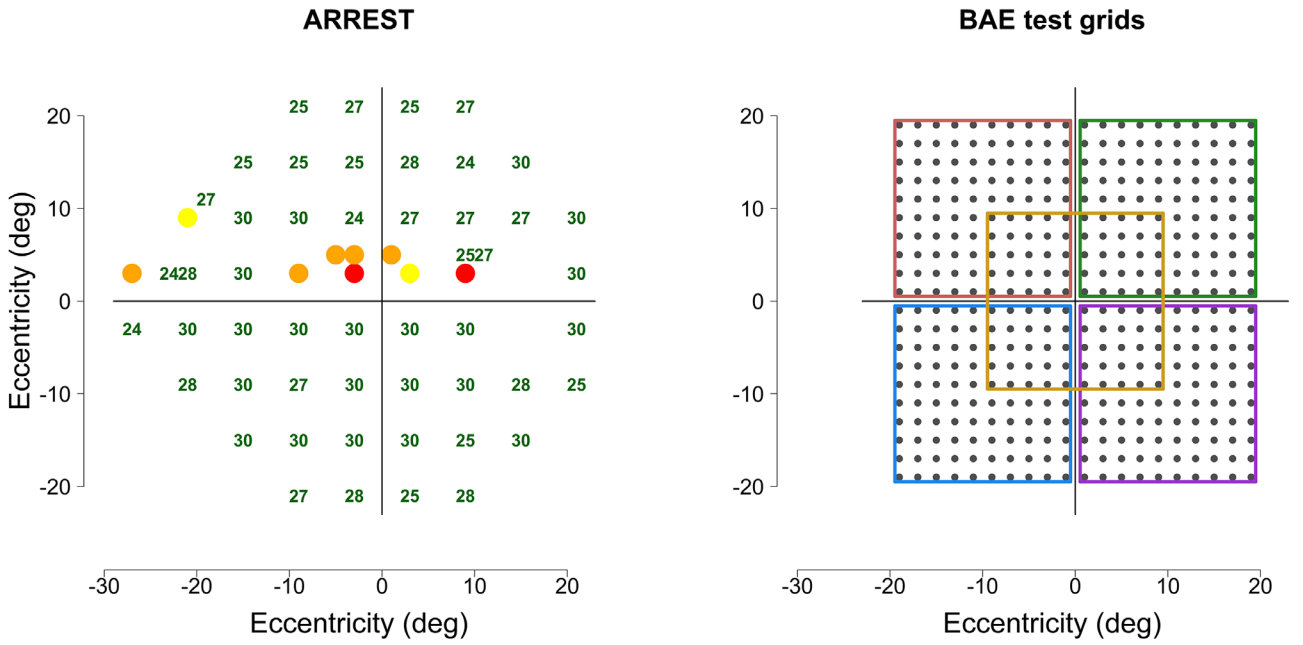

Figure 2. An example participant's visual field report for the ARREST approach (*left*). For the BAE (*right*), the macular grid was chosen as a region of interest for this participant (central grid denoted by the *yellow-brown square*). The *right panel* illustrates the quadrants or macular grid options for the assessment of BAE. Four different colored squares (*red*, *green*, *blue*, and *purple*) indicate each test quadrant.

perimeter has the ability to incorporate fundus tracking during the perimetric procedure. Fundus tracking was implemented for 13 of the 21 participants (but was erroneously turned off for the first 8 people). To use the same test locations in the second session, we registered the newly acquired compass fundus image to the baseline examination and generated new visual field coordinates to sit on the same locations of the retina using custom code (Python, OpenCV). To obtain the final BAE threshold at each of the 100 locations, the threshold values were converted to a linear scale (apostilbs), arithmetically averaged, and converted back to the dB scale[27,28] using the relationship $asb = 10^{(4 - dB/10)}$.

Mean Total Deviation

Although the main intent of ARREST is to add spatial detail to visual field assessment, global indices have utility in comparison to prior literature, for risk calculation, assessing visual field change, and other summary purposes.[6,18,29] In this study, we assessed the feasibility of calculating a global index for visual fields measured with the ARREST approach. MTD is a global visual field index that represents the average deviations of threshold from age matched normative values.[18] In order to determine normal reference values for age for our tests, our visually healthy participants performed cZEST and ARREST on the Octopus 900 perimeter. A test procedure was repeated after an adequate break if any abnormal visual field defect was detected that was not consistent with the other procedure. If a test procedure was repeated, the first visual field test was excluded from calculating the normal reference range (regardless of the presence or absence of the visual field defect in the second test).

Iwase et al.[30] has shown that perimetric sensitivity deteriorates after approximately 50 years of age by −0.11 dB per year, regardless of eccentricity in those with healthy vision.[30] In this study, we adjusted threshold values at each test location to 50 years of age using −0.11 dB loss per year and obtained the median of all threshold values as a normal reference value for each test location and the procedure.

Once the normative values had been computed for a participant aged 50 years as above, we computed total deviation (TD) values for the glaucoma dataset at each test location by subtracting normative values from the threshold values, after correcting for the age of the participant. For calculating TDs for the test locations other than standard 24-2 test grid in ARREST, normative values were interpolated using the natural neighbor method.[31,32]

As MTD is a spatial average, and not an average of thresholds at one location, it was calculated by arithmetically averaging TD values of all the test locations for cZEST and spatially weighted values for the ARREST approach. As ARREST does not

necessarily sample from a uniformly spaced grid, a simple average of TD values will be biased toward the TD values where there is a higher density of locations. To counteract this, TDs are weighted by the normalized area of their tile in a Voronoi mosaic of the field (created using the tripack package in R). In the "yellow" and "orange" locations of the ARREST field, the true threshold may lie between 0 dB and 16 dB, and so some value must be chosen at these locations as the most likely estimate of "threshold" in order to compute TD (T_Y and T_O respectively; see Fig. 1). For locations that are classified as "yellow" or "orange" for the first time, complete thresholding with ZEST has already been performed as a first step, followed by stimulus presentations of 16 dB and/or 0 dB (see Fig. 1). In this study, we set T_Y and T_O to be the mean of the final pdf after the initial ZEST and the extra presentations. If a location showed a change from "yellow" at the start of the test to "orange" at the end, it remains with a value of T_Y. The "red" locations were included in the MTD calculation as 0 dB. All of these values are shown in boxes at the bottom of Figure 1.

Data Analysis

The performance of cZEST and ARREST approach were compared using both the number of locations tested and the number of stimulus presentations required. The number of stimulus presentations and the MTDs were statistically compared using the paired *t*-test after checking that assumptions for using parametric tests for paired samples were met. Finally, to assess the validity of the classification of visual field locations with estimated sensitivity < 17 dB, threshold values from the cZEST, ARREST, and the BAE were classified into two bins (< 17 dB and ≥ 17 dB). Only the test locations that were common in both the ARREST and cZEST were included. Statistical analysis was performed using RStudio[33] (Integrated Development Environment for R version 4.0.0).[34]

Results

Number of Locations and Stimulus Presentations Comparison

The average number of locations tested by ARREST was 59 ± 4 compared with the standard 52 locations in the cZEST. In this study, the mean number of new locations added by the ARREST was 7 (range = 0 to 15). ARREST and cZEST used 259 ± 25 and 261 ± 25 stimulus presentations, respectively.

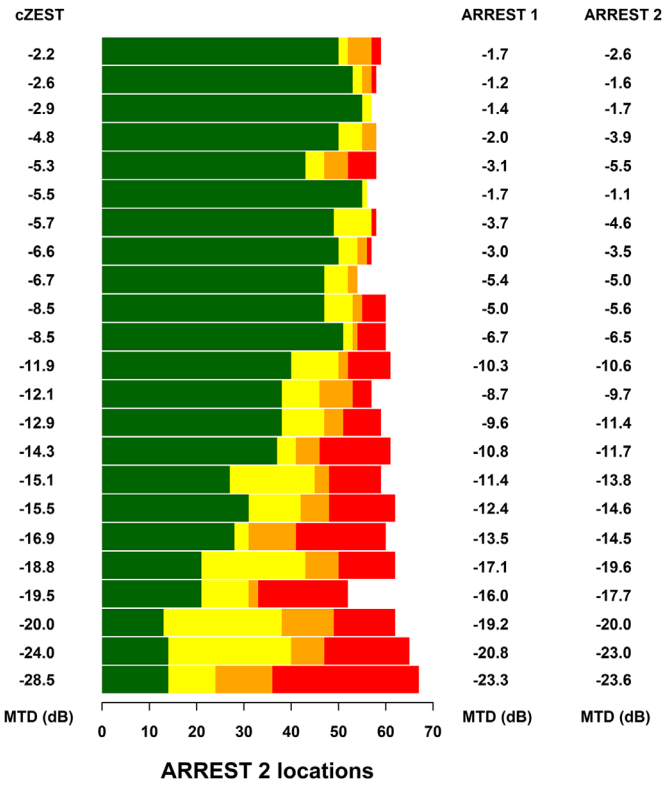

Figure 3. One row per participant showing the number of locations tested in ARREST 2 as the bar with corresponding MTDs of cZEST and ARREST fields, ordered by cZEST MTD. *Green* shows the number of locations with $>= 17$ dB, *yellow* the number of locations classified as 0 to 16 dB, *orange* is the number of locations classified as 0 dB, and *red* the number of locations classified as < 0 dB.

There was no statistically significant difference between ARREST and cZEST procedures for the total number of stimulus presentations ($t(22) = 0.29$, $P = 0.78$).

Comparison of MTDs of the ARREST and the cZEST

As noted above, cZEST and ARREST did not use identical ZEST procedures, hence we collected separate normative data for each. One effect of this difference in underlying ZEST procedures was that the threshold values for ARREST were limited to a maximum of 30.2 dB. In turn, this means that normative values were lower in some locations for ARREST than cZEST, hence we might expect the MTD for ARREST to be higher than cZEST. Figure 3 shows that this is indeed the case: on average, the MTD values for cZEST are 2.64 dB lower than ARREST 1 ($t(22) = -10.94$, $P < 0.001$; paired *t*-test). To observe the effect on MTD of adding additional spatial locations, a comparison was made between ARREST 1 and ARREST 2 (right hand side of Fig. 3). A paired *t*-test showed that the difference between ARREST 1 and ARREST 2 was

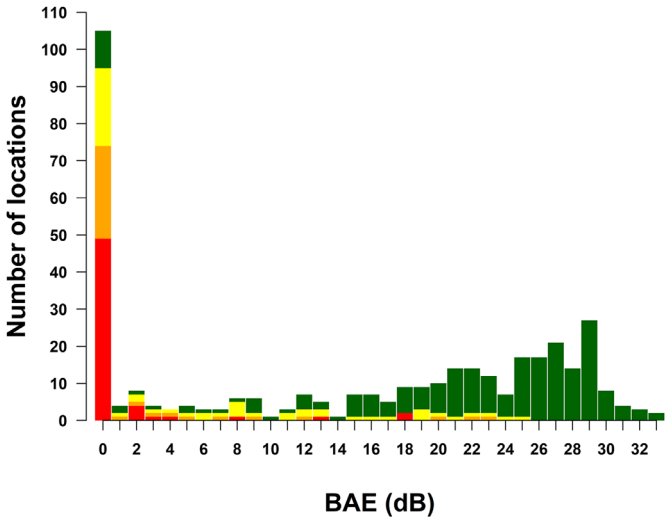

Figure 4. A count of locations with each BAE value in each of the four possible ARREST categories: *red*, *orange*, *yellow*, and *green* (as in Fig. 1).

statistically significantly different ($t(22) = 5.38$, $P < 0.001$, ARREST 1 mean −9.05 dB, SD = 6.78 dB; ARREST 2 mean −10.09 dB, SD = 7.16 dB) with the additional test points decreasing the MTD in most cases.

Classification of Highly Variable Locations (< 17 dB) by ARREST

The ARREST approach uses three or four stimulus presentation checks within a single test additional to the underlying ZEST to classify a location into the "yellow" bin (see Fig. 1). Once a location is classified as "yellow", it is no longer fully measured in future tests. Instead, these locations are presented only with stimuli of 0 dB and can progress to "orange" and "red" if 0 dB is unseen. Hence, misclassification of a location as "yellow" can mask any future changes in the location's sensitivity above 16 dB from visit to visit. Figure 4 shows the classification of locations in ARREST grouped by the BAE estimate for each location. Only spatial locations that occurred in both ARREST and BAE in the glaucoma dataset are shown ($n = 370$). Of the 190 locations that have a BAE above 16 dB, 10 are "yellow", 2 are "red", and 3 are "orange".

Use of ARREST Approach for Follow-Up

The ARREST approach is designed to add an increasing number of test locations as visual field loss severity increases, in order to improve tracking of progression (either by enabling visualization of the spatial spread of visual field loss, or by finding new "normal" locations that can be tracked for decreasing sensitivity over time). As a proof of concept, we retested 4 people from the glaucoma group with ARREST and cZEST approximately 1 year after the initial test date. The total number of locations tested in the 1-year follow-up visit (Fig. 5: right panels [ARREST 3]) by the ARREST approach for each of the participants (A, B, C, and D) were 66, 59, 64, and 68 which are 6, 5, 6, and 7 locations more than the initial visit (Fig. 5: middle panels [ARREST 2]) in this study. Note, these follow-up fields are not presented to illustrate progression per se, but to show how ARREST increases the spatial information available at each test, according to time savings that can be gained from the information in the prior field. The MTD for each of the test and follow-up fields is also provided.

Discussion

New locations added by the ARREST approach yield additional spatial information regarding visual field loss. Here, we demonstrate the feasibility of the approach in a cross-sectional study, hence the number of new locations added was for a single ARREST test. As demonstrated in Figure 5, if participants are followed in future visits, ARREST adds more locations at each test visit, providing the clinician with increased information about the spatial nature of the person's visual field.

In our previous computer simulation, we used a bimodal ZEST 24-2 procedure, which did not include a growth pattern for either the ARREST procedure or the ZEST reference. The application of growth patterns reduce test time but can increase bias in estimates.[23,35] In our previous simulation study, ARREST used 20% to 40% fewer stimulus presentations once visual field loss became relatively advanced compared to ZEST.[3] The ZEST procedure utilized was that implemented as the default ZEST in the OPI, which does not incorporate a growth pattern, and does not cap presentation counts per location, hence can be quite lengthy. In this study, to compare ARREST against a procedure designed to represent a clinical standard, we used cZEST where we did incorporate a growth pattern (for cZEST but not for ARREST), and capped the maximum number of presentations per location at 12. This resulted in the total number of stimulus presentations being similar for both the procedures with the key benefit of ARREST being the addition of new test locations.

Performing a single visual field test using ARREST took an average of 8 minutes and 23 seconds. In

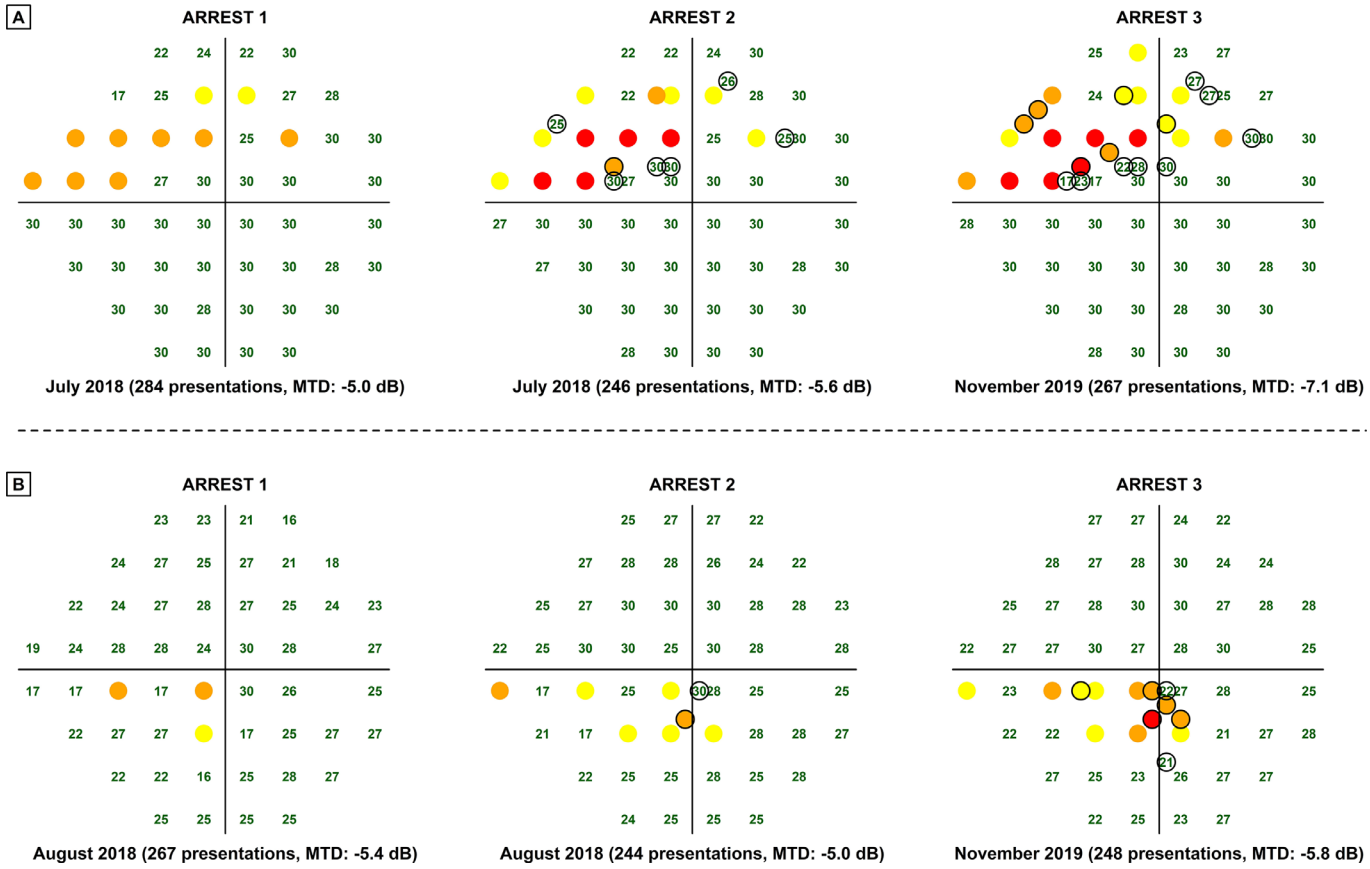

Figure 5. ARREST field reports of participants A (63 years old), B (69 years old), C (73 years old), and D (70 years old). ARREST 1 (*left*): Represents a regular clinical visual field test, however, the locations less than 17 dBs are binned and color coded according to the ARREST rules; ARREST 2 (*middle*): new locations added by ARREST (test locations in *black circle*) at the second test; ARREST 3 (*right*): Follow-up ARREST after 1 year adds additional test locations to the visual field while maintaining test presentation number.

this study, we did not try to optimize test duration, but instead a priori selected a desired number of 250 presentations. It is likely that the addition of a growth pattern, such as used in cZEST, would reduce the number of presentations required for ARREST, as would capping the number of permitted presentations per location. Adding a growth pattern to the underlying ZEST is simple (as in previous approaches), but care would need to be taken in subsequent tests to allow for any yellow or orange locations, or any additional locations in the field. The most beneficial method for spatial neighborhood logic when additional locations are added requires further exploration. Furthermore, there are various methods that are currently implemented in commercial perimeters, such as skipping false positives checks, not performing blind spot checks, or having dynamic response windows that significantly alter test duration.[36]

We designed this study to demonstrate that the ARREST approach can be used to gain more spatial information without increasing the test duration of the underlying test algorithm (in this case, cZEST). It should be noted that ARREST can be applied to any underlying procedure as a layer on top of whatever "bells and whistles" have been incorporated within the underlying commercial procedure to save test time. By way of an example, the ARREST approach of not fully thresholding locations with visual field sensitivity less than 17 dB, nor testing established perimetrically blind locations, could be added to already highly abbreviated test strategies, such as SITA-Faster[37] to allow individualized increased spatial resolution.

The key feature of ARREST is that it relies on information collected during preceding visual field tests from the same individual. Prior data is used to not only decide which locations require full assessment but also to determine where new added locations should be placed in the visual field. Furthermore, the determination of how many additional visual field points to add is also based on a conservative calculation of likely total test presentation numbers derived from the prior test result. For the implementation of ARREST explored herein, we simply use a relatively standard ZEST

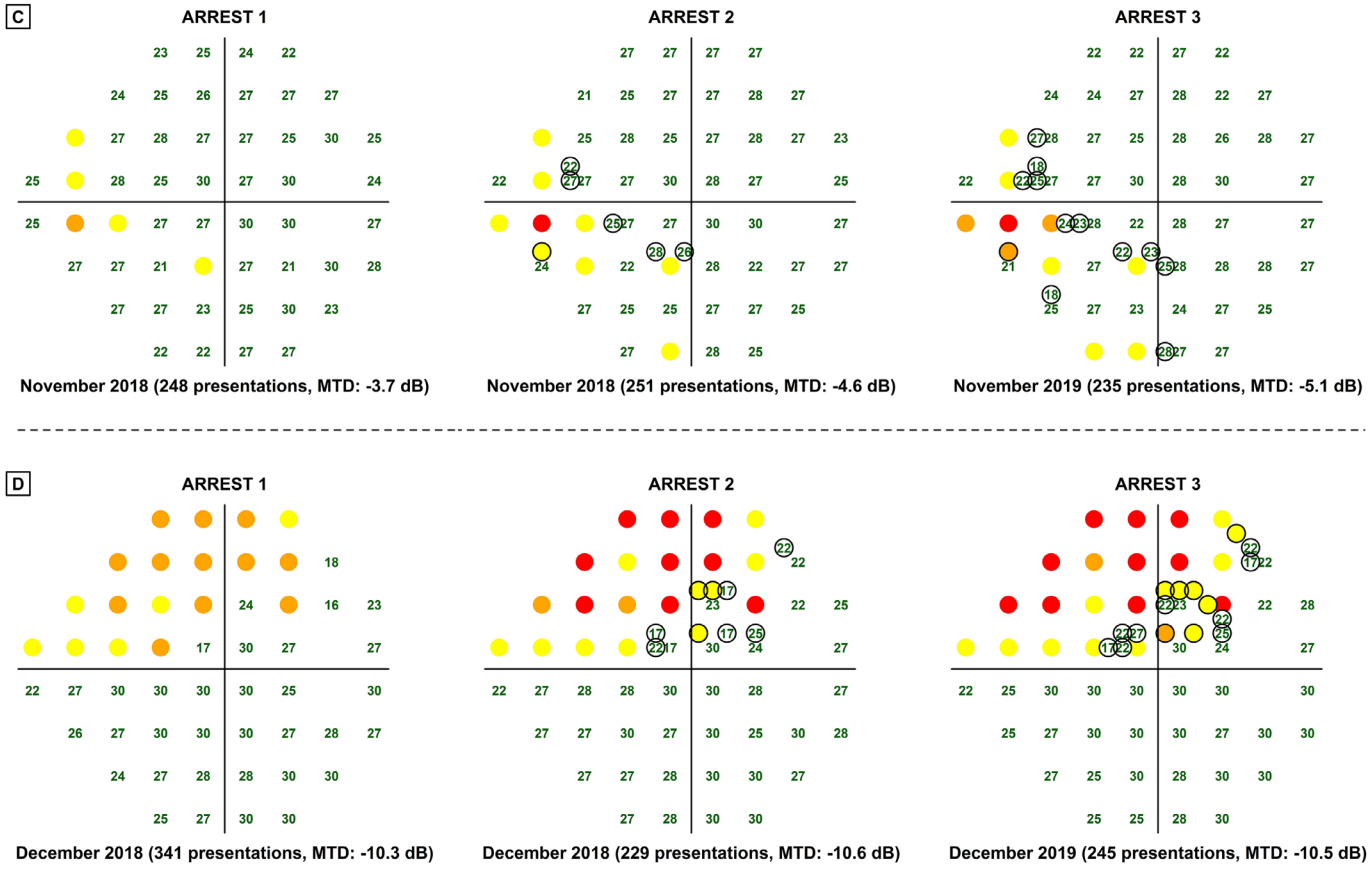

Figure 5. Continued

procedure as the underlying procedure. However, as noted above, ARREST can be applied to any test procedure, including those that use prior information to potentially speed up the test. For example, we have previously shown that information from prior visual fields can be used to bias the prior probability distribution for ZEST[38] or combined with suprathreshold testing strategies.[39] Depending on the preferred trade-off between test speed and spatial information gain, utilizing the prior visual field information in additionally sophisticated ways could result in either shorter tests, or increased spatial sampling. Further work is required to explore these trade-offs in detail, in particular, exploring the ability to both detect progression and also optimally assess areas of visual field important for tasks of daily living. Our current implementation of ARREST simply chooses the steepest gradient in the visual field as the region of interest for new test locations, however, this could also be biased by either prior information (for example, to add locations in an area different to the areas that have received additional locations in recent tests) or weighted somehow by relative importance to visual function, depending on what has been discovered at prior test visits.

The number of locations added by ARREST can vary markedly for eyes with advanced visual field damage for early tests in a testing sequence. For example, Figure 6 shows two participants with advanced field loss (cZEST MTD = −28.47 dB in participant E and −19.50 dB in participant F). ARREST added 15 additional locations for participant E (Fig. 6 top right panel) and 0 new locations for participant F (Fig. 6 lower right panel). The number of locations to be added was determined based on the first test in the series (left panels for participants E and F in Fig. 6). We did not collect additional tests for these two participants, however, we determined that if participants E and F were to undergo another visit, the ARREST approach would add 15 and 11 additional locations respectively while maintaining similar test duration as previous. Thus, our cross-sectional study represents seeing a new patient with existing field damage for their first two visual fields, but in a typical clinical scenario, where someone may be followed with regular visits over many years, the ARREST approach will add many additional locations compared with the results in this study.

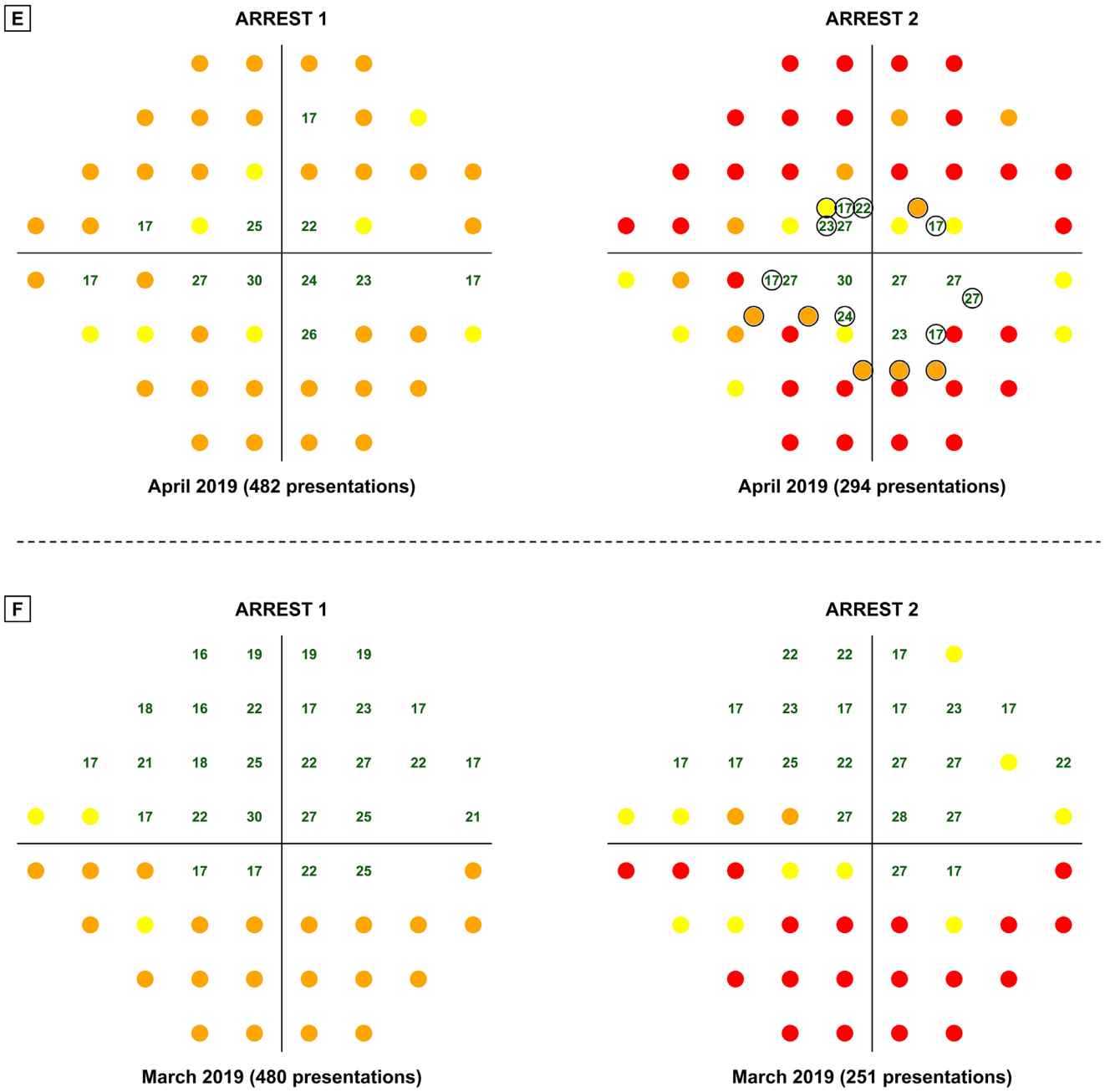

Figure 6. Examples of ARREST 1 and 2 fields (tested in a same day for this cross-sectional study) with 15 new location added for participant E (*top right panel*) and no locations added for participant F (*bottom right panel*). Newly added locations are denoted in the *black circle*.

Global visual field indices lack spatial information but are commonly used as outcome measures from clinical trials for disease staging and risk calculation.[6,29] In this study, we showed that ARREST fields can be interpreted to obtain global measures (MTD). This result implies that global indices for ARREST could be used across longitudinally acquired visual field data, however the best way to do this requires further study, particularly by simulation of the performance of ARREST on longitudinal data series. Because ARREST adds locations along the borders of scotoma, these new locations may mildly shift the MTD from one visit to the next through the addition of either more normal or damaged points. We account for this, in part, by spatially weighting our MTD index. The effect of adding locations on the MTD in our case examples is shown in Figure 3 where the MTD is compared between ARREST 1 and ARREST 2. As noted earlier, the difference in MTD for cZEST and ARREST 1 is driven by the difference in the underlying

ZEST procedures, rather than due to the ARREST framework per se. An alternate approach for calculating MTD that might be useful for progression could involve calculating the resultant MTD if all remaining possible visual field locations are interpolated at each visit from the existing measured locations. We note that computing global indices somewhat defeats the purpose of adding spatial information with ARREST, but is entirely feasible.

Previously, studies have shown that censoring threshold values below 20 dB had very little effect on glaucomatous progression assessment for either pointwise or censored MD trend analysis.[17,40] As concurrent change might occur in neighboring locations in moderate and advanced visual field loss, those studies have recommended shifting the lower end of the perimetric dynamic range to 15 to 19 dB[40] and to redesign newer perimetric algorithms to test more useful locations.[41] ARREST provides a framework for doing this, however, it does not completely eliminate testing below 15 to 19 dB. ARREST checks for visual field progression from 16 dB to 0 dB (yellow to orange or red) using spot checks. In highly advanced visual field loss, progression to perimetrically blind (i.e. from yellow to orange or red in ARREST) may provide meaningful information. For locations with sensitivities above 17 dB, standard progression criteria that are currently applied to longitudinal series of visual fields can be applied to data from ARREST (as described in our previous simulation study of progression analysis).[3]

Adding spatial locations complicates the analytical determination of visual field progression, however, our previous simulation study[3] explores solutions and demonstrates feasibility. Specifically, we have shown that event-based criteria that consolidates pointwise information work well. There are many event-based criteria that can be used,[42] however, all share four parameters: (1) a definition of baseline from which progression must occur, (2) a number of visual field points that must show a decrease in sensitivity from baseline, (3) the level of dB decrease that is considered important, and (4) the number of visits in a row where the decrease in sensitivity must be confirmed. Our previous simulation study demonstrates that this approach can be successfully implemented, with each new location requiring the relevant number of test repeats prior to consideration for progression.

Our current implementation of ARREST only differs from its underlying procedure when at least one location has sensitivity < 17 dB, and there is estimated to be enough presentation budget (aiming for approximately 250 presentations in total) to allow additional test locations to be added. One outcome of this approach is that entirely "clean" visual fields terminate in fewer presentations (approximately 210 on average). Hence, depending on the desired test duration, it may be possible to add test locations in situations of earlier visual field loss than the requisite 17 dB cutoff described here. For example, additional locations of known higher risk for glaucomatous damage in the macular area could be added,[43–45] or additional locations could be added in areas surrounding established scotomata where the sensitivity has not quite reached the 17 dB criterion. This may have particular benefit in exploring the spatial distribution of visual field loss close to fixation, in otherwise relatively normal fields. Note, whereas we have tested people with glaucoma in this study, ARREST is not a visual field procedure specific for glaucoma but can be used to explore the spatial nature of visual field defects in any condition. Further simulation work using empirical visual field series containing both 10-2 and 24-2 data may be helpful to explore the various trade-offs prior to recommending a more sophisticated approach.

In summary, here, we present a feasibility study of ARREST in a sample of people with glaucoma. The study demonstrates the feasibility of the ARREST approach applied to an underlying ZEST procedure; however, the approach can be applied to any extant visual field testing algorithm. We collected normative reference data here to determine TD for our laboratory algorithms implemented using the OPI, however, the ARREST framework applied to existing procedures would be able to utilize existing normative data. The ARREST framework has no features that are specifically designed to maximize performance for glaucoma. Hence, the principals involved in ARREST should be able to be applied for visual field testing of neurological defects, peripheral visual field defects, and for macular damage, however, empirical study will be necessary. The current iteration of ARREST tested herein would, however, require minor modification to the threshold checking procedure for situations where significant visual field recovery may be expected, stroke for example. As demonstrated in our previous simulations and by the case series herein, ARREST is designed to progressively improve the spatial description of the visual field with increasing numbers of follow-up visits, in the absence of increasing test duration.

Acknowledgments

The authors thank both anonymous reviewers both for their general comments, but specifically for helpful suggestions regarding the calculation of MTD.

Supported by Australian Research Council LP150100815.

Disclosure: **V. Muthusamy**, None; **A. Turpin**, Haag-Streit AG (R), Heidelberg Engineering GmBH (R), and CenterVue SpA (R, C); **M.J. Walland**, None; **B.N. Nguyen**, None; **A.M. McKendrick**, Haag-Streit AG (R), Heidelberg Engineering GmBH (R), and CenterVue SpA (R, C)

References

1. Bengtsson B, Heijl A. Evaluation of a new perimetric threshold strategy, SITA, in patients with manifest and suspect glaucoma. *Acta Ophthalmol Scand*. 1998;76(3):268–272.
2. Bengtsson B, Heijl A, Olsson J. Evaluation of a new threshold visual field strategy, SITA, in normal subjects. *Acta Ophthalmol Scand.* 1998;76(2):165–169.
3. Turpin A, Morgan WH, McKendrick AM. Improving spatial resolution and test times of visual field testing using ARREST. *Transl Vis Sci Technol*. 2018;7(5):35.
4. Leske MC, Heijl A, Hyman L, Bengtsson B. Early manifest glaucoma trial. *Ophthalmology*. 1999;106(11):2144–2153.
5. Garway-Heath DF, Lascaratos G, Bunce C, et al. The United Kingdom Glaucoma Treatment Study: a multicenter, randomized, placebo-controlled clinical trial: design and methodology. *Ophthalmology*. 2013;120(1):68–76.
6. De Moraes CG, Liebmann JM, Levin LA. Detection and measurement of clinically meaningful visual field progression in clinical trials for glaucoma. *Prog Retin Eye Res*. 2017;56:107–147.
7. Schiefer U, Malsam A, Flad M, et al. Evaluation of glaucomatous visual field loss with locally condensed grids using fundus-oriented perimetry (FOP). *Eur J Ophthalmol*. 2001;11:S57–S62.
8. Schiefer U, Flad M, Stumpp F, et al. Increased detection rate of glaucomatous visual field damage with locally condensed grids: a comparison between fundus-oriented perimetry and conventional visual field examination. *Arch Ophthalmol*. 2003;121(4):458–465.
9. Nevalainen J, Paetzold J, Papageorgiou E, et al. Specification of progression in glaucomatous visual field loss, applying locally condensed stimulus arrangements. *Graefes Arch Clin Exp Ophthalmol*. 2009;247(12):1659–1669.
10. Junoy Montolio FG, Wesselink C, Jansonius NM. Persistence, spatial distribution and implications for progression detection of blind parts of the visual field in glaucoma: a clinical cohort study. *PLoS One*. 2012;7(7):e41211.
11. Flammer J, Drance SM, Zulauf M. Differential light threshold. Short- and long-term fluctuation in patients with glaucoma, normal controls, and patients with suspected glaucoma. *Arch Ophthalmol*. 1984;102(5):704–706.
12. Heijl A, Lindgren G, Olsson J. Normal variability of static perimetric threshold values across the central visual field. *Arch Ophthalmol*. 1987;105(11):1544–1549.
13. Gilpin LB, Stewart WC, Hunt HH, Broom CD. Threshold variability using different Goldmann stimulus sizes. *Acta Ophthalmol*. 1990;68(6):674–676.
14. Henson DB, Chaudry S, Artes PH, Faragher EB, Ansons A. Response variability in the visual field: comparison of optic neuritis, glaucoma, ocular hypertension, and normal eyes. *Invest Ophthalmol Vis Sci*. 2000;41(2):417–421.
15. Wyatt HJ, Dul MW, Swanson WH. Variability of visual field measurements is correlated with the gradient of visual sensitivity. *Vision Res*. 2007;47(7):925–936.
16. Gardiner SK, Swanson WH, Goren D, Mansberger SL, Demirel S. Assessment of the reliability of standard automated perimetry in regions of glaucomatous damage. *Ophthalmology*. 2014;121(7):1359–1369.
17. Gardiner SK, Swanson WH, Demirel S. The effect of limiting the range of perimetric sensitivities on pointwise assessment of visual field progression in glaucoma. *Invest Ophthalmol Vis Sci*. 2016;57(1):288–294.
18. Artes PH, O'Leary N, Hutchison DM, et al. Properties of the statpac visual field index. *Invest Ophthalmol Vis Sci*. 2011;52(7):4030–4038.
19. Turpin A, Artes PH, McKendrick AM. The Open Perimetry Interface: an enabling tool for clinical visual psychophysics. *J Vis*. 2012;12(11):22.
20. King-Smith PE, Grigsby SS, Vingrys AJ, Benes SC, Supowit A. Efficient and unbiased modifications of the QUEST threshold method: theory, simulations, experimental evaluation and practical implementation. *Vision Res*. 1994;34(7):885–912.
21. Wall M, Doyle CK, Brito CF, Woodward KR, Johnson CA. A comparison of catch trial methods used in standard automated perimetry in glaucoma patients. *J Glaucoma*. 2008;17(8):626–630.
22. Turpin A, McKendrick AM, Johnson CA, Vingrys AJ. Development of efficient threshold strategies for frequency doubling technology perimetry using

computer simulation. *Invest Ophthalmol Vis Sci*. 2002;43(2):322–331.
23. Turpin A, McKendrick AM, Johnson CA, Vingrys AJ. Properties of perimetric threshold estimates from full threshold, ZEST, and SITA-like strategies, as determined by computer simulation. *Invest Ophthalmol Vis Sci*. 2003;44(11):4787–4795.
24. Anderson DR, Patella VM. *Automated static perimetry*. 2nd ed. Saint Louis, MO: Mosby; 1999.
25. McKendrick AM, Turpin A. Advantages of terminating Zippy Estimation by Sequential Testing (ZEST) with dynamic criteria for white-on-white perimetry. *Optom Vis Sci*. 2005;82(11):981–987.
26. Spry PG, Johnson CA, McKendrick AM, Turpin A. Measurement error of visual field tests in glaucoma. *Br J Ophthalmol*. 2003;87(1):107–112.
27. Hood DC, Anderson SC, Wall M, Kardon RH. Structure versus function in glaucoma: an application of a linear model. *Invest Ophthalmol Vis Sci*. 2007;48(8):3662–3668.
28. Hood DC, Kardon RH. A framework for comparing structural and functional measures of glaucomatous damage. *Prog Retin Eye Res*. 2007;26(6):688–710.
29. Smith SD, Katz J, Quigley HA. Analysis of progressive change in automated visual fields in glaucoma. *Invest Ophthalmol Vis Sci*. 1996;37(7):1419–1428.
30. Iwase A, Fujii M, Ohno Y, Araie M. Effects of myopia-associated factors and aging on visual field (VF) subfield sensitivities in normal eyes. *Invest Ophth Vis Sci*. 2019;60(9):4381.
31. Sibson R. A brief description of natural neighbour interpolation. In:Barnett V, ed. *Interpreting Multivariate Data*. Chichester: John Wiley; 1981:21–36.
32. Smith TB, Smith N, Weleber RG. Comparison of nonparametric methods for static visual field interpolation. *Med Biol Eng Comput*. 2017;55(1):117–126.
33. Team R. *RStudio: Integrated Development Environment for R*. Boston, MA: RStudio, Inc.; 2018.
34. Team RC. *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing; 2020.
35. Spenceley SE, Henson DB. Visual field test simulation and error in threshold estimation. *Br J Ophthalmol*. 1996;80(4):304–308.
36. Bengtsson B, Olsson J, Heijl A, Rootzen H. A new generation of algorithms for computerized threshold perimetry, *SITA*. *Acta Ophthalmol Scand*. 1997;75(4):368–375.
37. Heijl A, Patella VM, Chong LX, et al. A new SITA perimetric threshold testing algorithm: construction and a multicenter clinical study. *Am J Ophthalmol*. 2019;198:154–165.
38. Turpin A, McKendrick A. RE–ZEST: using patient rather than population information for visual field retest. *Invest Ophthalmol Vis Sci*. 2005;46(13):3731.
39. Turpin A, Jankovic D, McKendrick AM. Retesting visual fields: utilizing prior information to decrease test-retest variability in glaucoma. *Invest Ophthalmol Vis Sci*. 2007;48(4):1627–1634.
40. Pathak M, Demirel S, Gardiner SK. Reducing variability of perimetric global indices from eyes with progressive glaucoma by censoring unreliable sensitivity data. *Transl Vis Sci Technol*. 2017;6(4):11.
41. Wall M, Zamba GKD, Artes PH. The effective dynamic ranges for glaucomatous visual field progression with standard automated perimetry and stimulus sizes III and V. *Invest Ophthalmol Vis Sci*. 2018;59(1):439–445.
42. Vesti E, Johnson CA, Chauhan BC. Comparison of different methods for detecting glaucomatous visual field progression. *Invest Ophthalmol Vis Sci*. 2003;44(9):3873–3879.
43. Chen S, McKendrick AM, Turpin A. Choosing two points to add to the 24-2 pattern to better describe macular visual field damage due to glaucoma. *Br J Ophthalmol*. 2015;99(9):1236–1239.
44. Hood DC, Nguyen M, Ehrlich AC, et al. A test of a model of glaucomatous damage of the macula with high-density perimetry: implications for the locations of visual field test points. *Transl Vis Sci Technol*. 2014;3(3):5.
45. Phu J, Kalloniatis M. Ability of 24-2C and 24-2 grids to identify central visual field defects and structure-function concordance in glaucoma and suspects. *Am J Ophthalmol*. 2020;219:317–331.