



Detailed Evaluation of Chromatic Pupillometry and Full-Field Stimulus Testing to Assess Ultralow Vision in Retinitis Pigmentosa

Midori Yamamoto,^{1,*} Take Matsuyama, PhD,^{1,*} Tadao Maeda, MD, PhD,¹ Seiji Takagi, MD, PhD,² Naohiro Motozawa, MD, PhD,¹ Daiki Sakai, MD,^{1,3,4} Yasuhiko Hiramami, MD, PhD,^{1,3} Akiko Maeda, MD, PhD,¹ Yasuo Kurimoto, MD, PhD,^{1,3} Masayo Takahashi, MD, PhD,^{1,5} Michiko Mandai, MD, PhD¹

Purpose: Novel therapeutic options, such as regenerative medicine and gene therapy, are now emerging as viable treatment options for patients with severe visual impairments, such as retinitis pigmentosa (RP). Gradable assessment of patients' visual function is essential to consider treatment options and to evaluate treatment outcomes; however, evaluation of visual function in patients with advanced low vision is often challenging because of patients' poor and sometimes unpredictable responses. In this study, we attempted to accurately assess visual capabilities and disease stage in patients with RP with a visual acuity (VA) of ≤ 0.01 .

Design: Retrospective analysis of visual function indicators, including VA, retinal thickness, full-field stimulus testing (FST), and chromatic pupillometry.

Subjects: Overall, 43 patients (84 eyes) with advanced RP with a VA of ≤ 0.01 visited Kobe City Eye Hospital from 2019 to 2021.

Methods: Hierarchical (multilevel) Bayesian modeling was used to estimate individual eye's pupil response and FST threshold, taking into account the ambiguity and randomness often observed in patients with ultralow vision. Using the estimated ability obtained from each test, the correlation between each test and retinal thickness was further analyzed to make a comprehensive assessment of the data.

Main Outcome Measures: Visual acuity, retinal thickness, FST threshold, and pupil diameter change to different light stimuli.

Results: Full-field stimulus testing and pupillometry measurements were moderately correlated with VA but exhibited a wide range of values within the same VA groups. Full-field stimulus testing was not correlated with central retinal thickness at counting fingers/hand motion VA range and seemed to reflect overall remaining photoreceptor function, including peripheral retina. Pupillometry may be able to distinguish between different levels of inner retinal function.

Conclusions: The combination of pupillometry and FST allowed for graded evaluation of visual function within patients grouped in the same VA groups in patients with advanced RP with ultralow vision.

Financial Disclosure(s): Proprietary or commercial disclosure may be found after the references. *Ophthalmology Science* 2023;3:100328 © 2023 by the American Academy of Ophthalmology. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).



Supplemental material available at www.opthalmologyscience.org.

Retinitis pigmentosa (RP) is an inherited retinal degeneration with a reported prevalence of 1 in 4000 and is one of the leading causes of severe visual dysfunction worldwide.¹ In typical RP, the preceding loss of rod photoreceptor cells leads to night blindness and concentric visual field constriction, followed by a progressive loss of cone photoreceptor cells resulting in deterioration of central vision, sometimes leading to blindness.

As with many neurodegenerative diseases, there have been no established treatments to restore visual function once it is lost. However, in recent years, retinal cell-based

therapies and gene therapies have been used in clinical trials for severe retinal diseases. These new therapies conceptually allow for high expectations of not only slowing down disease progression but also restoring visual functions. Optogenetic intervention, which aims to restore light responsiveness by ectopic expression of light-activated proteins in inner retinal cells, is a novel gene therapy approach that has been applied to human patients with promising results.² Regenerative medicine is a replacement therapy for damaged or lost cells to regenerate visual function. We reported a proof-of-concept study for

pluripotent stem cell–derived retinal organoid transplantation in animal models of RP, in which we showed evidence for host-graft synapse formation, restoration of retinal ganglion cell (RGC) light responses after transplantation, and improvement in light-evoked behavior.^{3,4} Based on these proof-of-concept studies, we recently conducted a clinical trial for patients with RP using human-induced pluripotent stem cell–derived retinal organoids (jRCTa050200027).

Development and clinical application of these new treatments mandate an accurate and gradable assessment of patients' visual function because these target specific points of the visual pathway and disease stage. The most commonly used measure of visual function is the visual acuity (VA), which measures the ability to recognize small details with precision. Unfortunately, patients with severe visual dysfunction, for example, patients with a VA of < 0.01, who are primary candidates for these new treatments, are often grouped together into broad categories based on a simplified VA test: counting fingers (CF), hand motion (HM), light perception (LP), and no light perception (NLP). Each of these off-charts VA tests focus on different modalities of vision, such as the ability to sense movement or light. Crucially, the test only reflects patients' abilities under a predefined set of conditions, which are not rigorously standardized, and patients that show negative responses may exhibit positive responses under different conditions. Furthermore, patients with severe visual dysfunction may often experience light flashes (photopsia), which may further obfuscate patients' true visual capabilities.

A more detailed assessment of which part of the visual pathway has a residual function and to what extent will inform the best treatment options. If the therapeutic concept is to enhance photoreceptor function through gene therapy or other means, it is important to determine the degree of photoreceptor survival in the retina. For photoreceptor transplantation, it is crucial that secondary and tertiary neurons that transmit signals from photoreceptors are functioning or can become functional after photoreceptor transplantation. Optogenetics and artificial eye approaches would require a functioning neural pathway from RGCs to the visual cortex. A more accurate assessment of visual function and disease stage will help determine the indications for and methods of treatment. Additionally, the same methodology could be used to reliably determine the effectiveness of the treatment and assess the extent of improvement in visual functions, if any.

To assess visual function in patients with late-stage RP with low vision, a number of functional tests have been reported, including full-field stimulus testing (FST)^{5–7} and chromatic pupillometry.^{8–11} Full-field stimulus testing is a subjective test using full-field light stimuli, which can theoretically separate the cone and rod photoreceptor responses by using different wavelengths of the stimulus. Chromatic pupillometry is a noninvasive and objective method to study pupillary responses induced by cone and rod photoreceptors and melanopsin-containing intrinsically-photosensitive RGCs, which can be distinguished by the intensity and wavelength (color) of light stimuli.^{10,12,13} Light information for pupillary light reflex is relayed to

the brain through a pathway separate from the conventional visual pathway, with optic tract fibers terminating at the pretectal nucleus in the midbrain and not at the lateral geniculate nucleus of the thalamus.^{14–18} Selected stimulus conditions can produce pupil responses that reflect phototransduction primarily mediated by rods, cones, or melanopsin/intrinsically-photosensitive RGCs. We conducted a detailed analysis of FST and chromatic pupillometry on patients with advanced RP with ultralow vision. We then examined the correlation between the results of these tests and those of the conventional VA test as well as OCT measurements. Our results indicate that chromatic pupillometry and FST together with OCT can be used complementarily to more precisely evaluate visual function in patients with advanced RP and make better assessments on the estimated pathological status of patients.

Methods

Patients

This retrospective observational study protocol was approved by the institutional review board of Kobe City Medical Center General Hospital with the informed consent of the participants (approval number, E18006). The study was conducted in accordance with the terms of the Declaration of Helsinki. Written informed consent was obtained from all participants before data collection after an explanation of the nature and possible consequences of the study. We included patients with RP who visited the Kobe City Eye Hospital, Kobe, Japan, between 2019 and 2021. Inclusion criteria for patients were that the VA was equal to or lower than 20/2000, and those with other diseases that may affect visual function, such as optic nerve disease and uveitis, were excluded. Examinations were performed in 84 eyes of 43 subjects (Table 1), and all patients underwent a series of ophthalmic examinations that included the following: best-corrected VA, intraocular pressure, fundus examination, spectral-domain OCT, chromatic pupillometry, and FST. All measurements were performed by certified orthoptists and ophthalmologists (M.Y., T. Maeda, and N.M.) following standardized protocols.

VA Measurements

Best-corrected VA was tested using a standard Landolt C acuity chart. If the patient was unable to read the chart at 1-m distance, the following criteria were used to determine off-chart VA (CF, HM,

Table 1. Clinical Characteristic of the Patients*

	84 eyes of 43 patients
Patient age (yrs)	59.3 ± 9.96
Sex (female:male)	17:26
BCVA	
0.01	13 (15.5%)
CF	1 (1.2%)
HM	23 (27.4%)
LP	40 (47.6%)
NLP	7 (8.3%)

BCVA = best-corrected visual acuity; CF = counting fingers; HM = hand motion; LP = light perception; NLP = no light perception.

*Values are presented as number (%) or mean ± standard deviation.

LP, and NLP). The patient's VA is determined as CF if the patient can correctly count the number of fingers presented in front of their face. If the patient is unable to count fingers, the examiner waves their hand in front of the patient: the patient is considered HM if they can determine the direction of the HM. Lastly, if the patient is unable to detect HM, a torch is shone into the patient's pupil in the dark. The patient's VA is recorded as LP if they can sense the light, and NLP otherwise. The VAs were converted to logarithm of the minimum angle of resolution (logMAR) following the convention of Johnson et al¹⁹: a VA of 0.01 was assigned a logMAR VA of 2.0, a CF of 2.6, an HM of 2.9, an LP of 3.1, and an NLP of 3.4.

Spectral-domain OCT Evaluation

The spectral-domain OCT images were obtained using the Heidelberg Spectralis (Heidelberg Engineering, Spectralis). Total retinal thickness and thicknesses of the ganglion cell complex, inner nuclear layer, and photoreceptor layer (photo) were measured at 4 cardinal points 2000 μm from the fovea in horizontal and vertical scans. All measurements were performed using the "caliper" function of the Heidelberg instrument. The total retinal thickness was defined as the distance between the signal peak at the vitreoretinal interface (the internal limiting membrane) and the posterior boundary of the major signal peak that corresponds to the basal retinal pigment epithelium/Bruch's membrane complex. The ganglion cell complex was defined as the 3 innermost retinal layers: the nerve fiber layer, the ganglion cell layer, and the inner plexiform layer. The inner nuclear layer was defined as the distance between the basal inner plexiform layer and the outer plexiform layer. Lastly, the photoreceptor layer was defined as the region between the outer plexiform layer and the retinal pigment epithelium/Bruch's membrane complex.

Chromatic Pupillometry

Chromatic pupillometry and FST measurements were performed using an Espion system (Diagnosys LLC) with a ColorDome light stimulator (Diagnosys LLC) in a completely dark room. The protocol for chromatic pupillometry consisted of the following 4 steps. For each step, measurements were repeated 5 times. The first step (rod) consisted of a $-3 \log \text{cd/m}^2$ blue light stimulus after 10 minutes of dark adaptation. The second step (cone1) consisted of a $1 \log \text{cd/m}^2$ red light stimulus on 0.1cd/m^2 blue background light after 2 minutes of light adaptation. The third step (cone2) consisted of a $3 \log \text{cd/m}^2$ white stimulus light on 1cd/m^2 white background light after 2 minutes of light adaptation. Lastly, the fourth step (mela) consisted of a blue light stimulus of 150cd/m^2 without background light after 2 minutes of dark adaptation. Rod and cone responses are observed as transient responses immediately after light stimulation, whereas melanopsin response is observed as a sustained pupil response. Each test was performed on both eyes simultaneously, and changes in pupil diameter were recorded with an infrared camera in the ColorDome.

FST

Full-field stimulus testing was conducted with a 2-button alternative forced choice paradigm with an auditory cue before stimuli using the same equipment as that used for chromatic pupillometry (Espion system). After 45 minutes of dark adaptation after adding mydriatic eye drops in both eyes, blue (448 nm), white (590 nm), green (530 nm), and red (627 nm) light stimuli (4-ms flash) with photonically matched intensities were used to examine the rod and cone functions. The baseline of 0 decibels (dB) was defined as 0.01cd/m^2 for all stimulus colors. Subjects were tested monocularly, with the fellow eye shielded. The Profile System (Diagnosys LLC)

estimates the threshold using a 2-parameter Weibull function, taking into account false positives and false negatives.

Statistical Analyses

We used full Bayesian statistical inference with Markov Chain Monte Carlo sampling for statistical modeling using Rstan (Stan Development Team. 2017. RStan: the R interface to Stan. R package version 2.21.5. <http://mc-stan.org>). We employed Bayesian inference because of its flexibility in specifying the model structure, which was critical in this study, to account for some irregular features of the data. Posterior distributions, presented with mean and 95% intervals, represent the credible intervals for estimated parameters. Because most data consist of measurements gathered from both eyes of subjects, we incorporated terms for patient and eye biases to account for the hierarchical structure of the data. Using these and other relevant parameters, such as the color of stimulus light, we obtained posterior estimates from linear regression models. We used generic weakly informative priors for our Bayesian analysis to allow the data to primarily drive the inferences without imposing strong assumptions about the parameter values. Note that posterior distributions do not represent a simple pooling (average) of data for a particular set of predictor combinations, rather, they represent the effect of predictors while considering the data as a whole. Detailed explanation of the statistical models, including prior assumptions, as well as the Stan code for the models, are provided in a GitHub repository (<https://github.com/matsutakehoyo/Clinical-Data-Analysis>).

Results

Robust Hierarchical Modeling Allows for FST Evaluation of Ultralow Vision Patients with Credible Intervals

Full-field stimulus testing measurements were analyzed using a multilevel (hierarchical) logistic model, assuming that the ability to respond to the light stimulus is influenced by patient/eye biases and stimulus color, in addition to the stimulus strength (see Fig S1, available at www.ophtalmologyscience.org, for details).

We saw a very clear effect for the color of the stimulus light, with patients generally being more sensitive to blue light than to red light (Fig 2A), as is expected from the spectral characteristics of photopic (cones) and scotopic (rods) systems. A key aspect of the model is the inclusion of the "guessing" parameter (α), which accounts for data points that do not conform well with logistic regression. Although such data may often be disregarded as unreliable, these may actually reflect features of patients' visual function because patients with RP can experience spontaneous flashes of light, which would obfuscate FST results. Therefore, we assumed that responses may originate from the normal process of subjects responding to light stimuli or a second random process. The

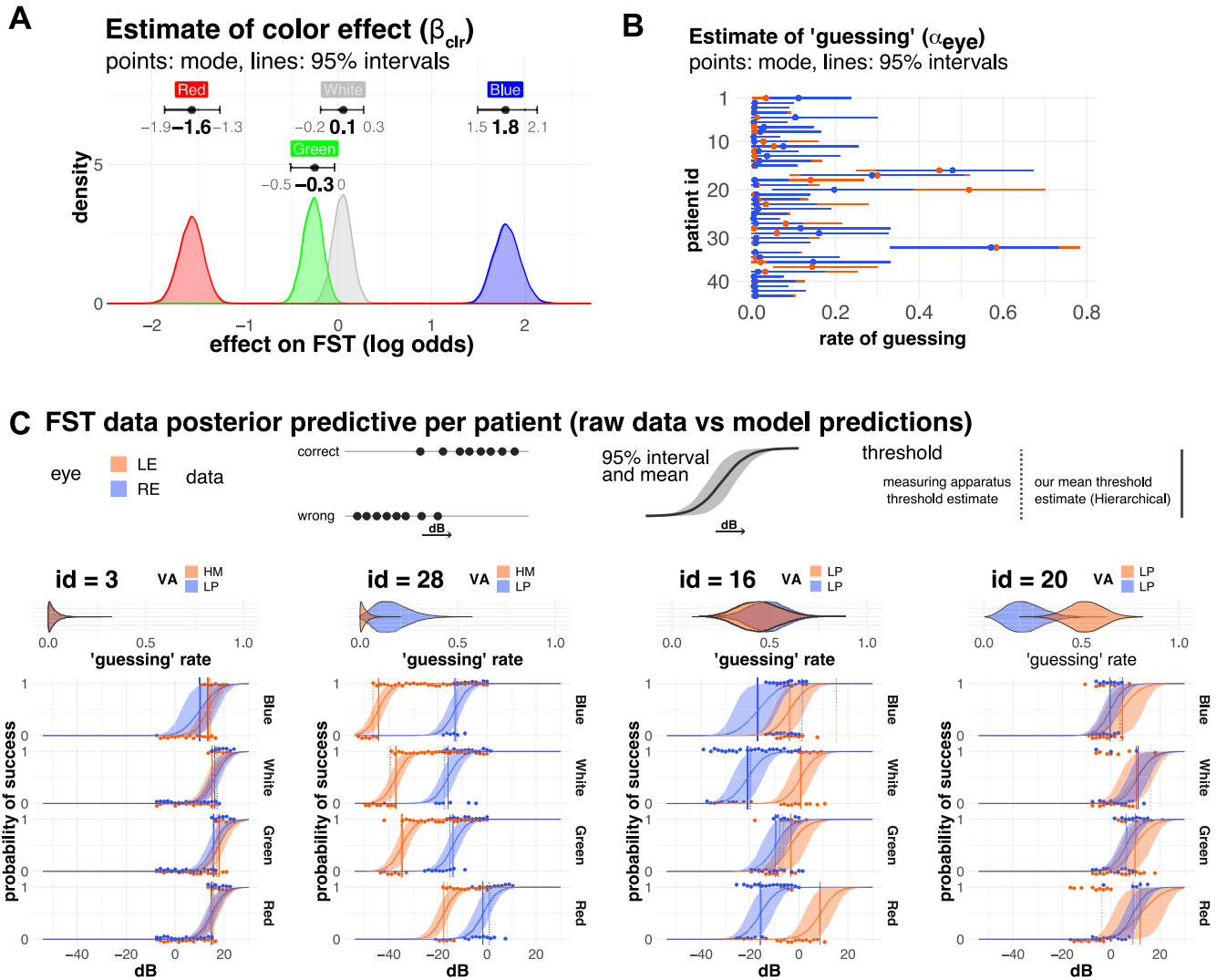


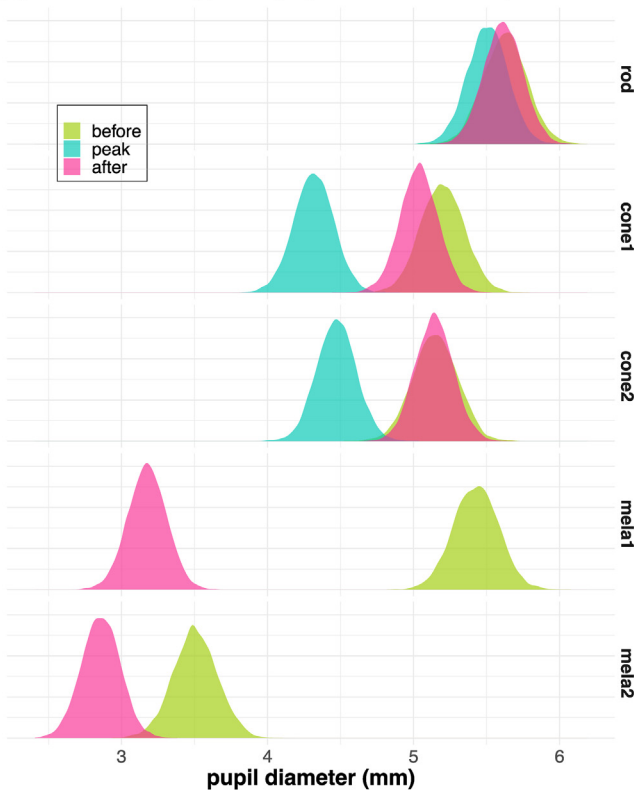
Figure 2. Analysis of full-field stimulus testing (FST). **A** and **B.** Posterior estimates of FST model. **A,** Estimates for the effect of light color (β_{clr}), with subjects being more sensitive to blue light than red light (Red < Green, White < Blue). **B,** Although most of the subjects had a relatively small "guessing" estimate (α_{eye}) with the mode and 95% interval below 0.1 (= 10% guessing), there were notable exceptions with very large guessing values (mode > 0.2). **C,** Comparison of raw data and model estimates for representative data. Points show the raw data for a series of experiments for a patient. Note that points are shown slightly offset from their true values (0 or 1) to avoid right eye (RE) and left eye (LE) points overlapping. Model estimates are shown with a solid curved line (mean) and ribbon (95% intervals). "Guessing" estimates are shown as violin plots on top of the FST results. The vertical lines show the FST threshold values (decibel at which probability of success is 0.5). The solid line represents the mean threshold value of our estimates (hierarchical), whereas the dotted line shows the threshold estimated by the measuring apparatus (threshold estimated individually). Color indicates LE (orange) and RE (blue). dB = decibels; HM = hand motion; ID = identification number; LP = light perception; VA = visual acuity.

"guessing" parameter represents the fraction of responses originating from the random process.

Figure 2B shows the estimated guessing for each patient's eye (α_{eye}). Although most of the eyes showed small "guessing" parameter (mode, < 0.05), there were notable exceptions where subjects seemed to be "guessing" more than half of the time (Fig 2B). Figure 2C shows representative cases with posterior predictions from our model alongside data measurements with the "guessing" rate presented on top of each panel. Typically, a subject's probability of correctly identifying the light stimulus decreases as the light intensity (x_i) decreases, as

described by the S-shaped curve of the logistic function. Vertical lines show the threshold value (the light intensity at which the probability of success = 0.5). Sensitivity to color generally decreases from blue to red, as previously stated, although the amount of shift varies from patient to patient and even from right eye (OD) to left eye (OS) (identification number [ID] = 3 and ID = 28 for example). In samples with larger "guessing," subject's responses were seemingly random, with several measurements falling outside the S-shaped curve of logistic regression (patient ID = 16, 20, and 28 OD). Although the differences in threshold values between our

A Estimated typical pupil diameter response



B Pupillometry data posterior predictive check

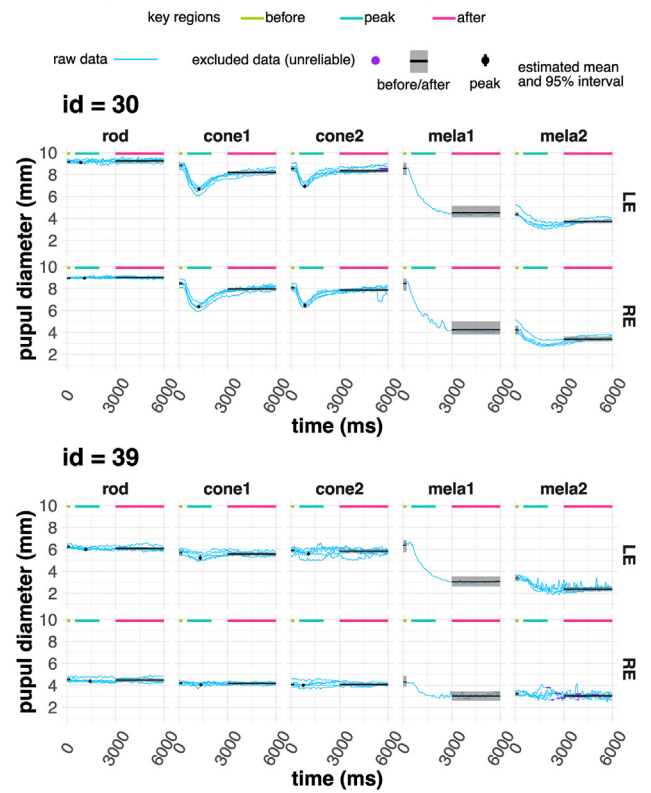


Figure 4. Analysis of chromatic pupillometry. **A**, Estimated typical pupil diameter change ($\beta_0 + \beta_{clr}$). **B**, Examples of raw data and model estimates highlighting typical and noteworthy cases. Light blue traces show the individual pupillometry measurements. Purple dots overlaid on the blue traces (identification number [id] = 39) show data points that were excluded from the analysis because they consisted of regions with constant values (standard deviation = 0) and, therefore, likely represented regions where pupil detection failed. Rod, cone1, and cone2 measurements consisted of 5 traces (repeats), whereas the 5 traces of mela light stimulus were separated as mela1 (repeat 1) and mela2 (repeats 2–5) because we noticed that there was a large difference between mela1 and mela2. The black and gray shaded area shows the estimated mean and 95% confidence interval for before, peak, and after pupil diameter. LE = left eye; RE = right eye.

estimates (solid line) and the conventional estimates made by the measurement instrument (dotted line) are practically identical in most of the cases, there are notable cases where they differ significantly (ID = 16 blue and ID = 20 red). There are 2 main differences between our estimates and the conventional estimates. The first one is the inclusion of the aforementioned “guessing” parameter (α_{eye}), which can account for outliers. The second one is that our estimates take into account the hierarchical nature of the data. We estimate the overall trend of the study cohort (β_0) as well as biases for each patient/eye (β_{pat} and β_{eye}) and the overall effect of color (β_{clr}), as well as deviations from these overall tend for particular combinations of eye/color ($\beta_{eye \times clr}$; see Figure S1 for details; available at www.opthalmologyscience.org). The dotted line represents the conventional threshold estimate obtained by fitting a 2-parameter Weibull function. This value is obtained from a particular set of data points, the measurements for a particular sample (eye) at a particular color in isolation (no pooling). In contrast, our estimates take into account data from all conditions and from all samples simultaneously (hierarchical partial pooling). An

important consequence of this is that our estimates are more skeptical of measurements that fall outside the group means. For example, the conventional threshold estimate (dotted line, no pooling) would indicate that an ID=16/OD and an ID=20/OS have much higher thresholds for blue light than for red light, which is physiologically unlikely because the response to blue stimuli could be driven by rod and cones, whereas responses to red stimuli are driven by cones. In our hierarchical Bayes analysis, on the other hand, the model incorporates all available data to determine that the overall color effect (β_{clr}) is strong and that the magnitude of patient and eye biases (β_{pat} and β_{eye}) is generally smaller than the effect of color (β_{clr}), resulting in a more physiologically reasonable estimate without any assumptions other than the data structure. The large “guessing” is also an important factor here because some of the observations are seemingly random. Because these estimates are not strongly supported by the data, compatibility intervals encompass a wider range of values. These estimates are nonetheless useful because they narrow the reasonable and credible range of values. Similarly, using the available information, we were able to make inferences even if data were entirely missing for a

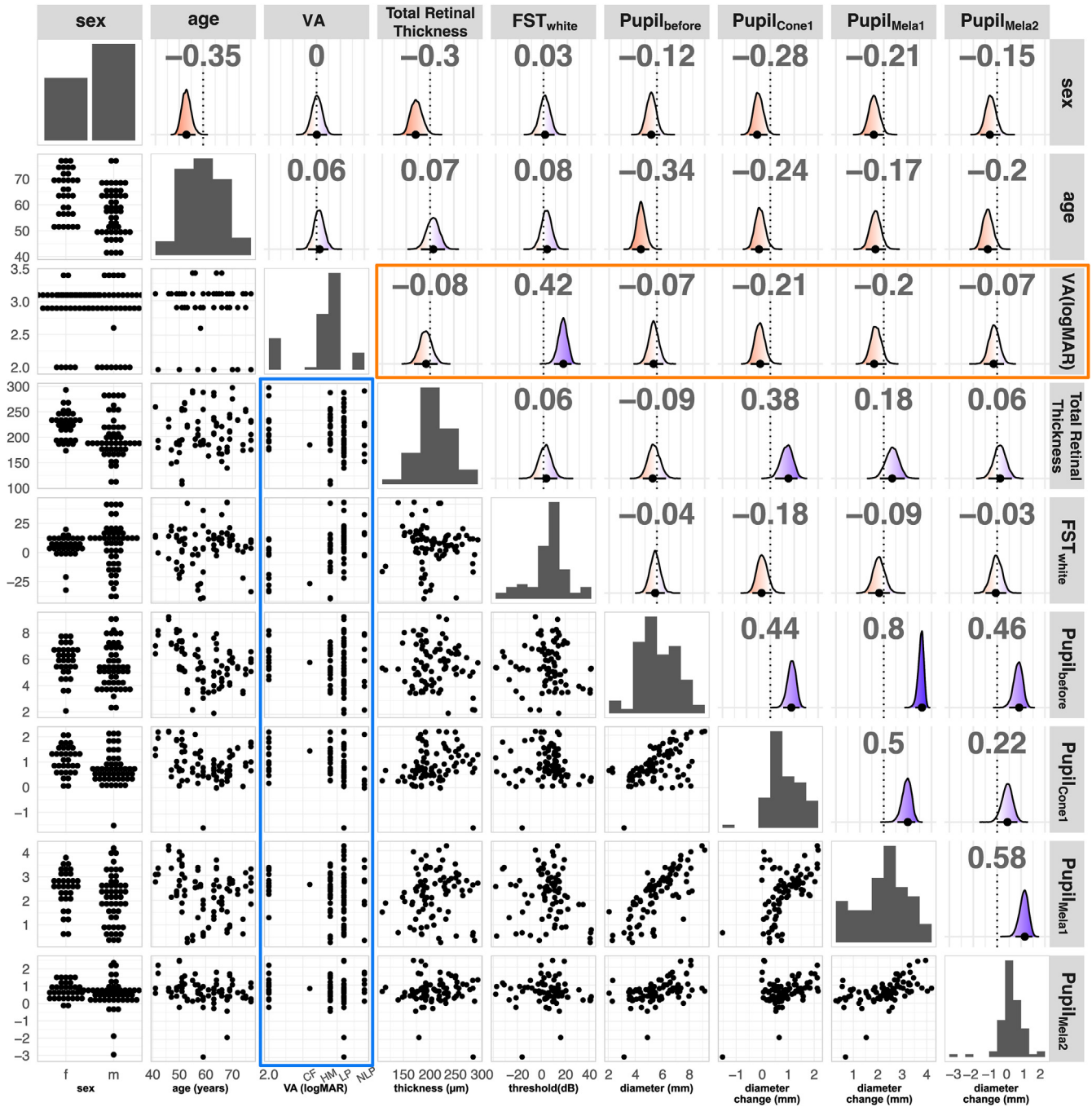


Figure 5. Pairwise relationship of dataset. We plotted (from left to right and top to bottom) sex, age, visual acuity (VA, logarithm of the minimum angle of resolution [logMAR]), total retinal thickness (μm), full-field stimulus testing (FST) (decibels, white light stimulation), pupil diameter before light stimulation ($Pupil_{before}$, mm), pupil change to cone1 stimuli ($Pupil_{cone1}$, mm), pupil change to mela1 stimuli ($Pupil_{mela1}$, mm), and pupil change to mela2 stimuli ($Pupil_{mela2}$, mm). The diagonal shows distribution of each variable (bar chart for categorical data and histogram for continuous data). Lower triangular panels show the pairwise relationship of features. The blue rectangle highlights the relationship of VA (logMAR) to other measurements. Upper triangular panels show the estimated correlation coefficient, estimated from a multivariate model, with density plots showing the distribution of estimated correlation coefficients. Red indicates negative correlation ($P < 0$), and blue indicates positive correlation ($P > 0$). The mean (black point) and 95% confidence interval (black bar) are indicated on the bottom of the density plot. Numbers on top indicate the mean, and the dotted line denotes the zero line. The orange rectangle highlights the correlation coefficients of features with VA.

particular color (ID = 3 blue for example) or in cases where no correct responses were obtained within the tested range. Although we employed the hierarchical Bayesian approach,

we acknowledge that the advantages of hierarchical Bayes models, such as handling missing data, can also be achieved with other statistical methods. Overall, we believe that our

model is able to capture and evaluate the uncertainty in measurements, giving us more reliable estimates with appropriate confidence intervals.

Analysis and Characterization of Pupil Responses in Eyes with Advanced RP

Pupillometry measurements were analyzed using a multi-level (hierarchical) multivariate model as illustrated in [Figure S3A](#) (available at www.opthalmologyscience.org). Measurements consist of time series of the pupil diameter changes to different stimulus (stimulus is applied at $t = 200$ ms). For each stimulation condition (rod, cone1, cone2, and mela), there were 5 repeats. Although repeated measurements seemed to reliably reproduce pupil diameter changes for rod, cone1, and cone2 stimulations, this was not the case for the mela stimulation. In most cases, the first repeat elicited a large reduction in pupil diameter, while repeats 2 to 4 exhibited a more attenuated response. Therefore, we separated the response to the mela stimulation into mela1 and mela2 to differentiate between the first and the remaining repeats. Although the data are a time series, we focused on 3 key regions, before, peak, and after light stimulation, to simplify the analysis. We extracted the pupil diameters from these regions for each measurement and estimated their values. We then used multivariate regression to estimate the overall trend of the study cohort (β_0) as well as biases for each patient/eye (β_{pat} and β_{eye}) and the overall effect of color (β_{clr} , the different stimulation conditions aiming at rod, cone, or melanopsin stimulations) as well as deviations from these overall trends for particular combinations of eye/color ($\beta_{eye \times clr}$), similar to the FST analysis (see [Figure S3](#)). [Figure 4A](#) shows the expected typical response ($\beta_0 + \beta_{clr}$). Note that peak was not defined for mela1 and mela2 because we did not expect or observe a sharp peak for these. Rod, cone1, cone2, and mela1 measurements had a baseline (before) of approximately 5.5 mm (rod, 5.6 mm; cone1/cone2, 5.2 mm). There was very little or no pupil change to the rod stimulation, as is expected in patients with RP. The rod/cone (cone1 and cone2) stimulation induced a small change, with pupils transiently constricting to approximately 4.3 to 4.5 mm and immediately reverting back to baseline values. On the other hand, there was a large constriction of the pupil for mela1 (3.2 mm), which only partially reverted (3.5 mm), resulting in very small pupil diameter change for mela2. [Figure 4B](#) highlights some examples with raw data alongside model posterior predictions ($\beta_0 + \beta_{clr} + \beta_{pat} + \beta_{eye} + \beta_{eye \times clr}$), showing that the model reasonably describes the data. [Figure 4B](#) ID = 30 shows a typical response, with an absent rod response, a small peak response for cone1 and cone2 stimuli, a large mela1 response (repeat #1), and a very small mela2 response (repeat #2–5). [Figure 4B](#) ID = 39 shows an example where rod and cone responses are absent, although there is a robust mela1 response. Moreover, notably, there is a significant difference between OS and OD. Additionally, we observed several noteworthy cases, such as those with completely albescent or gradually decreasing mela2

responses, and cases exhibiting oscillations with quick changes in the order of a few millimeters and 1 to 3 Hz. Although these phenomena may also be useful in diagnosing visual function, further investigation is warranted to better understand the significance of these findings.

FST and Pupil Response is Largely Uncorrelated to VA and Retinal Thickness

Having analyzed FST and pupil response separately, we next focused our attention on how these correlated to each other (see [Table S2](#), available at www.opthalmologyscience.org, for a list of all the values). [Figure 5](#) shows the pair plot for some representative features. We selected sex, age, VA (logMAR), total retinal thickness within macular area (2 mm from central fovea), FST (white), pupil diameter before stimulation ($Pupil_{before}$), and pupil change to cone1 ($Pupil_{cone1}$), mela1 ($Pupil_{mela1}$), and mela2 ($Pupil_{mela2}$) stimulation for this analysis. We used the threshold for white light stimulation for FST measurements as representative data for FST measurements because correlation between FST threshold for different stimulus color (red, green, white, and blue) was extremely high ($P > 0.9$; data not shown) and using any color would have resulted in practically identical results. Note the wide distribution of values for total retinal thickness, FST, and pupil diameters within the same VA group ([Fig 5](#), blue rectangle), indicating that individuals within the same VA group may exhibit a wide range of visual functions and retinal thicknesses. The correlation analysis showed that VA was not significantly correlated to retinal thickness or pupil responses, but moderately correlated to FST ([Fig 5](#), orange rectangle; $\rho_{FST/VA} = 0.42$). Because the central macular cone area is generally considered last to degenerate in eyes with RP, we initially expected macular retinal thickness to reasonably correlate with measures of visual function (FST or pupil response). However, we only found a moderate correlation with cone pupil response ($Pupil_{cone1}$) in these eyes with low vision. Retinal thickness also seemed to be correlated to subject's sex, with female subjects exhibiting a slightly thicker retina. This seems to be consistent with a previous study that reported differences in the progression of central vision function according to sex.²⁰

[Figures 6](#) and [7](#) are per-patient/eye summaries of collected data highlighting some characteristic cases. As we noted above, FST and pupil response does not correlate with VA or retinal thickness. The eye in [Figure 6](#) shows that the retinal thickness and the retinal layers are relatively well preserved, with clearly remaining outer nuclear layer in the fovea with VA of HM. However, FST was poor and with a rather high guessing rate. We further looked into the Goldman visual field test to find that the eye had a very constricted central vision. In contrast, [Figure 7](#) shows the eye with a very thin retina in the macular area with VA of CF. However, the eye exhibited a relatively good FST score with a clear shift between blue and red lights, indicating rod functionality. Pupil response to rod stimuli

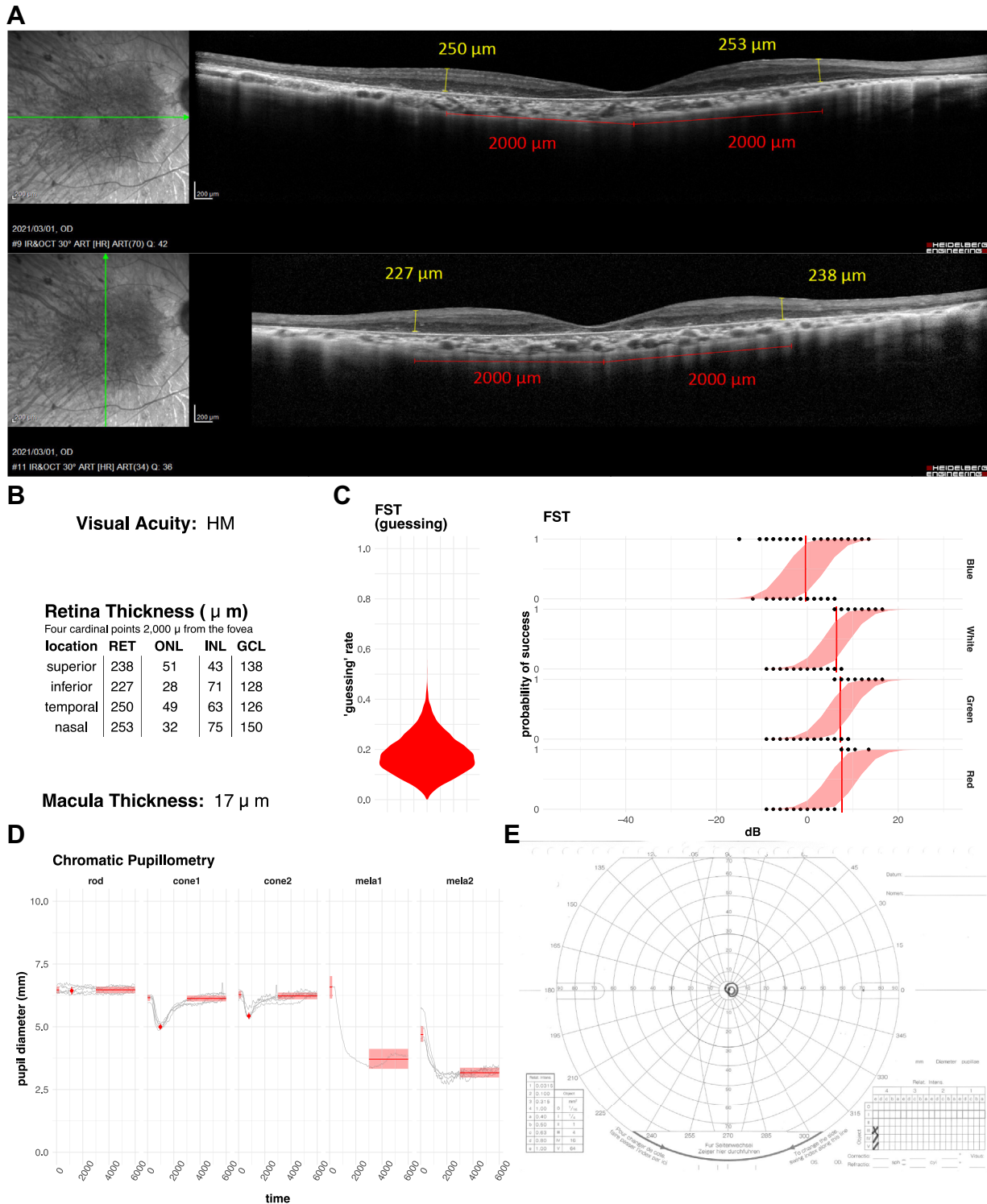


Figure 6. Representative case of an eye with a thick retina and poor full-field stimulus testing (FST). Vertical and horizontal OCT (A) scans show the moderately remaining outer nuclear layer (ONL) of the fovea with a relatively preserved retinal structure. The average retinal thickness (RET) near the macula (B) was 242 μm , which is equivalent to that of a normal subject. The eye had been diagnosed as hand motion (HM). All of the pupillay responses (D) except for rod response were clearly present. On the other hand, FST sensitivities (C) were relatively poor with threshold values over 0 decibels (dB). The result of the Goldman visual field test (E) shows that a very constricted visual field remained in the center. RET = retina; ONL = outer nuclear layer; INL = inner nuclear layer; GCL = ganglion cell layer

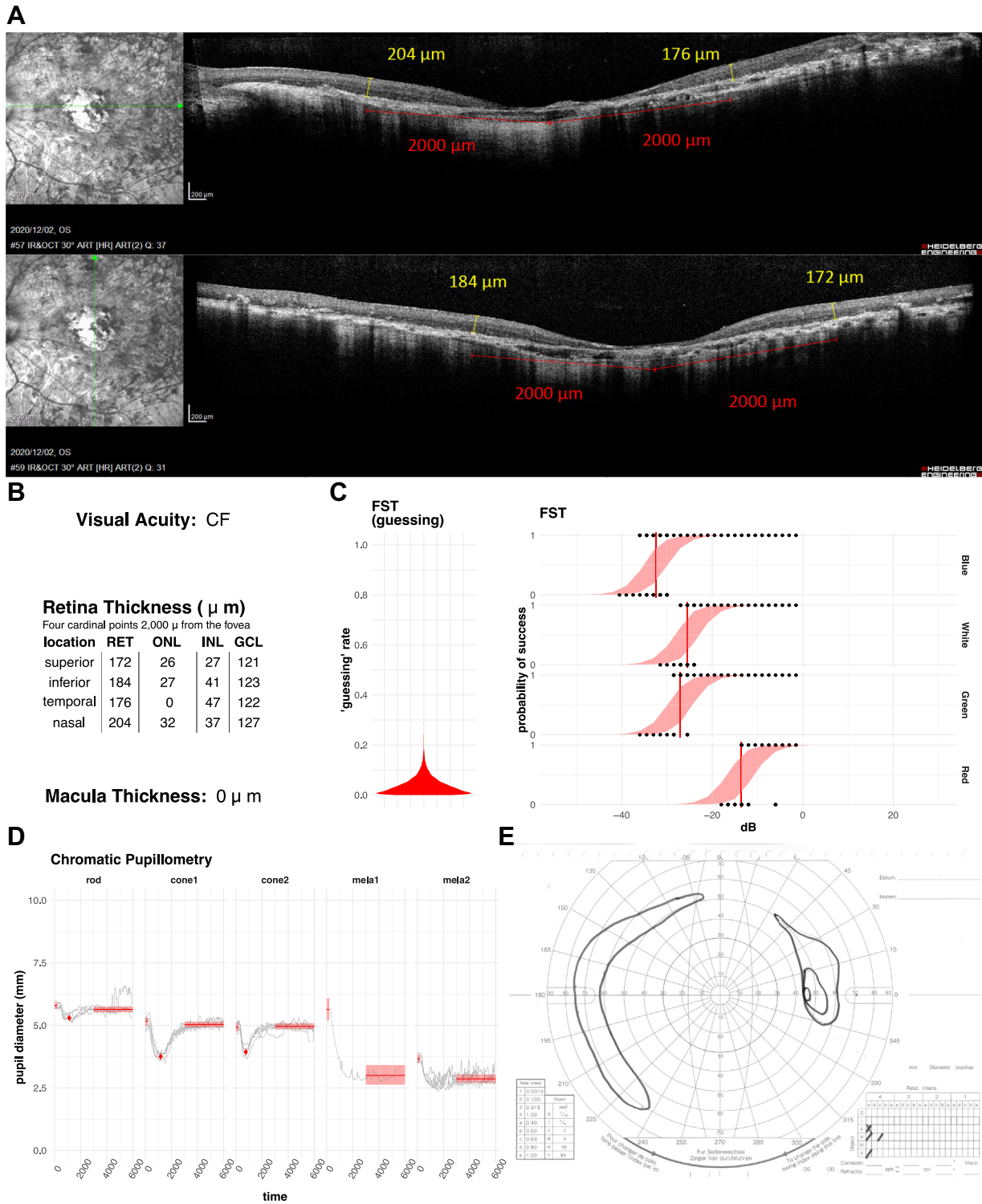


Figure 7. Representative case with a thin retina but good visual acuity and full-field stimulus testing (FST) sensitivity. The average retinal thickness (RET) (B) was 184 μm, and OCT images (A) revealed severe photoreceptor cell loss at the fovea. Visual acuity was counting fingers (CF), the highest category of this study, and the FST sensitivities (C) were similar to those of normal eyes, with a high sensitivity to blue light. The response to chromatic pupillometry (D) was also good with rod response, which was rare in our study cohort and the Goldman visual field test (E) showed a relatively large residual field of view in the periphery. RET = retina; ONL = outer nuclear layer; INL = inner nuclear layer; GCL = ganglion cell layer

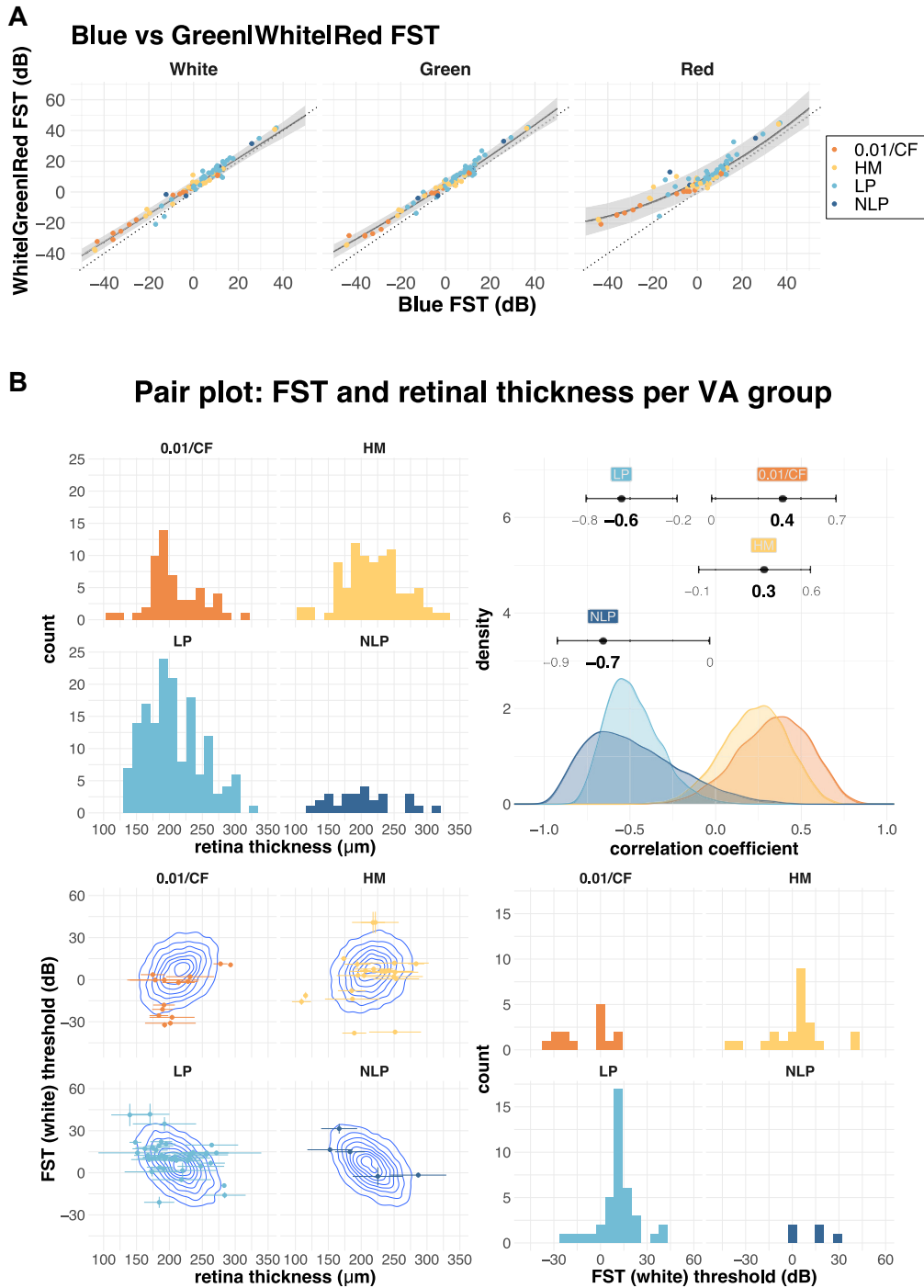


Figure 8. **A**, Relationship between full-field stimulus testing (FST) values for blue, white, green and red light. The blue FST value was plotted against white, green, and red FST values. The dotted diagonal line has an intercept of 0 and a slope of 1. Points above the dotted line indicate a higher FST value than blue light, and points below the dotted line indicate a lower FST value. The solid line shows the mean, and the shaded area shows the 95% compatibility interval of quadratic regression for each of the colors. **B**, Relationship between retinal thickness and FST (white) stratified by 4 visual acuity (VA) groups. The upper left panel shows the distribution of retinal thickness, and the lower right panel shows the distribution of FST for white light. The lower left panel shows the distribution of the data (points) overlaid with estimated correlation (density plot), with horizontal error bars showing the standard deviation of retinal thickness measurement (average of 4 points), and vertical error bars showing the estimated error (standard deviation) from the FST analysis. The upper right panel shows the distribution of posterior estimates for the correlation between FST and retinal thickness (error bars show 95% confidence interval). CF = counting fingers; dB = decibels; HM = hand motion; LP = light perception; NLP = no light perception.

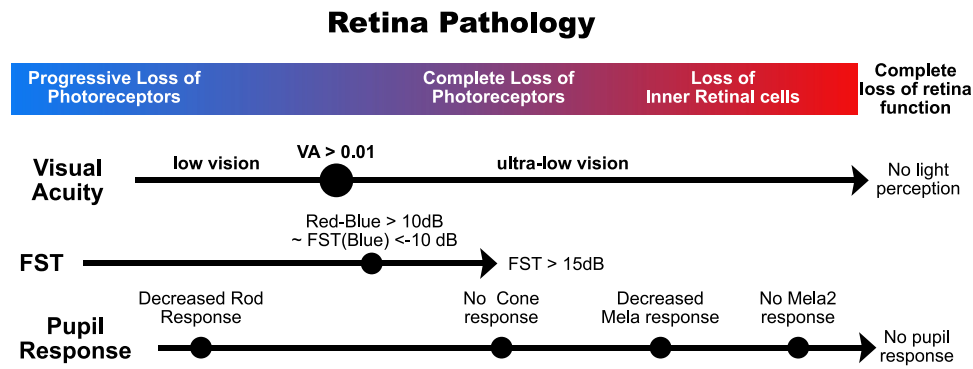


Figure 9. Full-field stimulus testing (FST) and pupillometry can be used complementary to visual acuity (VA) and OCT observations to more precisely estimate pathological stage. As photoreceptors degenerate, FST threshold increases. We used a red–blue FST of > 10 decibels (dB) as a reference point, indicating advanced rod loss. According to our analysis in Figure 6A, this is equivalent to a blue FST of < –10 dB. The loss of photoreceptor cells in the central macula as well as the peripheral area leads to > 15 dB FST test results and the loss of pupillary cone response. Further degeneration and loss of retinal inner cells result in a decrease in pupillary melanopsin responses.

was weakly present, supporting the presence of rod function in addition to cone function, which was rare in our patients with RP. The Goldman visual field test showed the presence of peripheral vision. Together, these results indicate that both cone and rod functions were preserved in spite of the very thin central retina.

Discussion

In this study, we employed a hierarchical Bayesian approach, which allowed us to handle large outliers and explore individual patient data. However, our study, being more data-driven than hypothesis-driven because of its exploratory nature and the lack of existing literature, might lead to overfitting or biased estimates with a small number of patients/eyes. We encourage further research with larger sample sizes and a priori hypotheses to validate and expand our findings. For instance, our study suggests a speculative connection between the guessing parameter and spontaneous flashes of light in patients with RP, which warrants additional investigation to better understand the underlying mechanisms and validate this hypothesis.

Our study highlights the importance of exercising caution when dealing with patient data, where assumptions for normal subjects may not necessarily hold and conventional approaches may not be appropriate. It is common or standard practice to repeat measurements to average and increase the quality of measurements, assuming that responses are reproducible. This was clearly not the case in our pupillometry measurements because melanopsin responses attenuated significantly after the initial measurement. Furthermore, in our FST analysis, we were able to recover more reasonable estimates from noisy data by taking into account potential random responses and incorporating all the available data. The loss of function associated with disease progression not only results in change of a particular metric, such as FST threshold or pupil diameter change, but also results in change in the reliability with which we can measure it. This highlights the importance of accounting for uncertainty in measurements and appropriately evaluating it

when dealing with patient data. These estimates with uncertainty intervals may aid clinicians to further narrow and diagnose the pathological background. Birch et al²¹ performed an FST study in patients with inherited retinal disease with a VA of ≥ 0.25 and reported that, in some cases, the threshold difference between blue and red stimuli was > 10 dB. The decrease in the difference between blue and red thresholds is generally thought to reflect the loss of rod function.^{5–7} Although our study cohort consisted of patients with more advanced retinal degeneration than that in the study by Birch et al,²¹ we were nonetheless able to detect differences between blue and red FST (18 eyes [21%]; [0.01 or CF {0.01/CF}: 6 eyes, HM: 6 eyes, LP]: 5 eyes, NLP: 1 eye). Furthermore, we noted that eyes that had low threshold values tended to have a larger shift between blue and red light than eyes that had high threshold values (e.g., compare ID = 28 and ID = 3 in Fig 2C, note that higher threshold values indicate lower probability to respond to the light stimulus). Birch et al²¹ have reported a similar relationship between white threshold and the difference between blue and red thresholds. We expand on this notion by modeling the relationship between blue threshold and other stimuli (green, white, and red) with a quadratic curve (Figure 8A). Although green and white values largely mirror blue threshold values in all regions, red threshold values show a larger deviation from blue values at lower thresholds. This is consistent with the notion that reduced sensitivity to blue light represents advanced rod degeneration. Thus, although FST measurements for white or green colors are highly correlated, blue and red FST values provide important diagnostic information in late-stage RP. Moreover, our data indicate that threshold values for white, green, and red light can reasonably be predicted from the blue light stimulus alone, or vice versa.

OCT measurements, such as retinal thickness, are widely and commonly used to diagnose retinal health, with well-organized thicker retinas are generally presumed to correlate with more visual function. Because central cones are often preserved in RP, we measured retinal thickness within the macular area. However, we found no correlation

between retinal thickness and indicators of visual function, such as VA, FST, or pupil response, in our cohort of advanced retinal degeneration with a VA of ≤ 0.01 (Fig 5). To further explore the relationship between OCT measurements and visual function, we stratified the data into 4 groups according to subject's VA (0.01/CF, HM, LP, and NLP). Unexpectedly, we found a reversal of correlation between retinal thickness and FST values (Fig 8). Patients with a thicker retina tended to have higher FST values (lower sensitivity for LP) in 0.01/CF and HM, whereas this trend was reversed in LP and NLP. Moreover, although Figure 8 shows the analysis using the average of 4 points for the entire retinal thickness (2000 μm from the fovea), the same trend was observed at each of the quadrant position individually (temporal, nasal, superior, and inferior) and for different layers (photoreceptor layer, inner nuclear layer, and ganglion cell layer), indicating that this is a global effect possibly affecting the entire macula. Although human visual function largely depends on central vision, example cases shown in Figures 6 and 7 suggest that in the eyes with advanced degeneration, remaining central visual function may be complemented by peripheral vision, explaining the apparent discrepancy between the retinal thickness around the macula and overall visual function. It is of note that the FST and pupillometry measurements seemed sensitive enough to suggest the presence of remaining rod and cone function in the peripheral vision in the eye with HM (Fig 5). These results highlight that, together, FST, pupillometry, and OCT measurements may help in understanding the pathological status of retinal degeneration, allowing for a more graded and

comprehensive assessment than the conventional VA categories of 0.01, CF, HM, LP, and NLP in patients with ultralow vision.

With the possibility of clinical application of novel treatments for visual restoration on the horizon, such as gene therapy, cell-based treatments, optogenetics, and prosthetic devices, a more detailed assessment of patients' visual potency and the state of the retina is urgently needed. Full-field stimulus testing is very sensitive and has been used to detect changes before and after treatment in a clinical study of gene therapy for patients with Leber congenital amaurosis.^{7,22,23} Indeed, in the current study, we obtained positive responses (defined as a threshold below + 15.0 dB) from almost half of the eyes with NLP, highlighting the sensitivity of this approach. In the pupillary test, the 2 cone stimuli resulted in practically identical results and seem to reflect the remaining photoreceptor pathway and RGC activities. Positive FST and cone pupillary test results indicate a functional retinal circuitry even in the eyes with poor central vision. In these eyes, gene therapy or cell-based approaches may help enhance visual responses. On the other hand, eyes that only have a pupillary response to melanopsin stimuli may better benefit from direct stimulation of RGCs, such as artificial eye or optogenetic treatment, rather than cell therapy. Lastly, eyes that completely lack FST and pupil responses may benefit from cortical interventions (Fig 9).

These emerging therapies for visual restoration bring hope for patients with severe visual impairments. The current study highlights the importance of developing diagnostic tools for patients with ultralow vision, bringing these treatments one step closer to clinical application.

Footnotes and Disclosures

Originally received: September 16, 2022.

Final revision: April 30, 2023.

Accepted: May 2, 2023.

Available online: May 8, 2023. Manuscript no. XOPS-D-22-00198R3.

¹ Department of Ophthalmology, Kobe City Eye Hospital, Kobe, Japan.

² Department of Ophthalmology, Toho University Graduate School of Medicine, Tokyo, Japan.

³ Department of Ophthalmology, Kobe City Medical Center General Hospital, Kobe, Japan.

⁴ Department of Surgery, Division of Ophthalmology, Kobe University Graduate School of Medicine, Kobe, Japan.

⁵ Vision Care Inc., Kobe, Japan.

*M.Y. and T.M. contributed equally to this work and are cofirst authors.

Disclosures:

All authors have completed and submitted the ICMJE disclosure form.

The authors have made the following disclosures: M.Y.: Grant – Government (during the conduct of the study) and Astellas, Santen, Sumitomo Pharma, TOMEY, NIKON, Vision Care, AMO. HOYA. Senju, Alcon, and Novartis (outside the submitted work).

T. Matsuyama: Government (during the conduct of the study) and Vision Care (outside the submitted work).

T. Maeda: Grant – Government (during the conduct of the study) and Astellas, Santen, Sumitomo Pharma, TOMEY, NIKON, Vision Care, AMO. HOYA. Senju, Alcon, and Novartis (outside the submitted work).

S.T.: Grant – Government (during the conduct of the study).

N.M.: Grant – Government (during the conduct of the study) and Astellas, Santen, Sumitomo Pharma, TOMEY, NIKON, Vision Care, AMO. HOYA. Senju, Alcon, and Novartis (outside the submitted work).

D.S.: Grant – Government (during the conduct of the study) and Astellas, Santen, Sumitomo Pharma, TOMEY, NIKON, Vision Care, AMO. HOYA. Senju, Alcon, and Novartis (outside the submitted work).

Y.H.: Grant – Government (during the conduct of the study) and Astellas, Santen, Sumitomo Pharma, TOMEY, NIKON, Vision Care, AMO. HOYA. Senju, Alcon, Novartis, Bayer, and Kowa (outside the submitted work).

A.M.: Grant – Government (during the conduct of the study) and Astellas, Santen, Sumitomo Pharma, TOMEY, NIKON, Vision Care, AMO. HOYA. Senju, Alcon, and Novartis (outside the submitted work).

Y.K.: Grant – Government (during the conduct of the study) and Astellas, Santen, Sumitomo Pharma, TOMEY, NIKON, Vision Care, AMO. HOYA. Senju, Alcon, Novartis, Kowa, Bayer, Pfizer, and VIATRIS (outside the submitted work).

M.T.: Grant – Government (during the conduct of the study) and Santen, Sumitomo Pharma, TOMEY, HealiOS, Nichirei Biosciences, Hitachi, Sysmex, MEIJI Seika, and SHIMADZU (outside the submitted work).

M.M.: Grant – Government (during the conduct of the study) and Astellas, Santen, Sumitomo Pharma, TOMEY, NIKON, Vision Care, AMO. HOYA. Senju, Alcon, Novartis, Otsuka, Bayer, and HealiOS (outside the submitted work).

Supported by Japan Agency for Medical Research and Development under grant number JP20bm0204002.

Presented at the 74th Annual Congress of Japan Clinical Ophthalmology, October 14-17, 2020.

HUMAN SUBJECTS: Human subjects were included in this study. This retrospective observational study protocol was approved by the institutional review board of Kobe City Medical Center General Hospital, with the informed consent of the participants (approval number, E18006). The study was conducted in accordance with the terms of the Declaration of Helsinki.

No animal subjects were used in this study.

Author Contributions:

Conception and design: Yamamoto, Tadao Maeda, Motozawa, Mandai.

Data collection: Yamamoto, Tadao Maeda, Motozawa.

Analysis and interpretation: Yamamoto, Matsuyama, Takagi, Sakai, Mandai.

Obtained funding: Takahashi, Mandai

Overall responsibility: Yamamoto, Matsuyama, Tadao Maeda, Takagi, Motozawa, Sakai, Hirami, Akiko Maeda, Kurimoto, Takahashi, Mandai.

Abbreviations and Acronyms:

CF = counting fingers; **FST** = full-field stimulus testing; **HM** = hand motion; **ID** = identification number; **logMAR** = logarithm of the minimum angle of resolution; **LP** = light perception; **NLP** = no light perception; **OD** = right eye; **OS** = left eye; **RGC** = retinal ganglion cell; **RP** = retinitis pigmentosa; **VA** = visual acuity.

Keywords:

Chromatic pupillometry, FST, Retinitis pigmentosa.

Correspondence:

Takeshi Matsuyama and Michiko Mandai, Kobe City Eye Hospital, 2-1-8 Minamimachi, Minatojima, Chuo-ku, Kobe, Hyogo 650-0047, Japan.
E-mail: matsutakehoyo@gmail.com; michiko_mandai@kcho.jp.

References

- Hartong DT, Berson EL, Dryja TP. Retinitis pigmentosa. *Lancet*. 2006;368:1795–1809.
- Sahel JA, Boulanger-Scemama E, Pagot C, et al. Partial recovery of visual function in a blind patient after optogenetic therapy. *Nat Med*. 2021;27:1223–1229.
- Assawachananont J, Mandai M, Okamoto S, et al. Transplantation of embryonic and induced pluripotent stem cell-derived 3D retinal sheets into retinal degenerative mice. *Stem Cell Rep*. 2014;2:662–674.
- Mandai M, Fujii M, Hashiguchi T, et al. iPSC-derived retina transplants improve vision in rd1 end-stage retinal-degeneration mice. *Stem Cell Rep*. 2017;8:69–83.
- Roman AJ, Schwartz SB, Aleman TS, et al. Quantifying rod photoreceptor-mediated vision in retinal degenerations: dark-adapted thresholds as outcome measures. *Exp Eye Res*. 2005;80:259–272.
- Roman AJ, Cideciyan AV, Aleman TS, Jacobson SG. Full-field stimulus testing (FST) to quantify visual perception in severely blind candidates for treatment trials. *Physiol Meas*. 2007;28:N51–N56.
- Roman AJ, Cideciyan AV, Wu V, et al. Full-field stimulus testing: role in the clinic and as an outcome measure in clinical trials of severe childhood retinal disease. *Prog Retin Eye Res*. 2022;87:101000.
- Kelbsch C, Maeda F, Lisowska J, et al. Analysis of retinal function using chromatic pupillography in retinitis pigmentosa and the relationship to electrically evoked phosphene thresholds. *Acta Ophthalmol*. 2017;95:e261–e269.
- Park JC, Moura AL, Raza AS, et al. Toward a clinical protocol for assessing rod, cone, and melanopsin contributions to the human pupil response. *Invest Ophthalmol Vis Sci*. 2011;52:6624–6635.
- Kardon R, Anderson SC, Damarjian TG, et al. Chromatic pupillometry in patients with retinitis pigmentosa. *Ophthalmology*. 2011;118:376–381.
- Lisowska J, Lisowski L, Kelbsch C, et al. Development of a chromatic pupillography protocol for the first gene therapy trial in patients with CNGA3-linked achromatopsia. *Invest Ophthalmol Vis Sci*. 2017;58:1274–1282.
- Rukmini AV, Milea D, Gooley JJ. Chromatic pupillometry methods for assessing photoreceptor health in retinal and optic nerve diseases. *Front Neurol*. 2019;10:76.
- Lorenz B, Strohmayer E, Zahn S, et al. Chromatic pupillography and RPE65 deficiency. *Invest Ophthalmol Vis Sci*. 2011;52:3891.
- Hattar S, Liao HW, Takao M, et al. Melanopsin-containing retinal ganglion cells: architecture, projections, and intrinsic photosensitivity. *Science*. 2002;295:1065–1070.
- Morin LP, Allen CN. The circadian visual system, 2005. *Brain Res Rev*. 2006;51:1–60.
- Güler AD, Ecker JL, Lall GS, et al. Melanopsin cells are the principal conduits for rod–cone input to non-image-forming vision. *Nature*. 2008;453:102–105.
- Czeisler CA, Shanahan TL, Klerman EB, et al. Suppression of melatonin secretion in some blind patients by exposure to bright light. *N Engl J Med*. 1995;332:6–11.
- Morin LP. The circadian visual system. *Brain Res Rev*. 1994;19:102–127.
- Johnson LN, Guy ME, Krohel GB, Madsen RW. Levodopa may improve vision loss in recent-onset, nonarteritic anterior ischemic optic neuropathy. *Ophthalmology*. 2000;107:521–526.
- Ito N, Miura G, Shiko Y, et al. Progression rate of visual function and affecting factors at different stages of retinitis pigmentosa. *Biomed Res Int*. 2022;2022:7204954.
- Birch DG, Cheng P, Duncan JL, et al. The RUSH2A study: best-corrected visual acuity, full-field electroretinography amplitudes, and full-field stimulus thresholds at baseline. *Transl Vis Sci Technol*. 2020;9:9.
- Russell S, Bennett J, Wellman JA, et al. Efficacy and safety of voretigene neparvovec (AAV2-hRPE65v2) in patients with RPE65-mediated inherited retinal dystrophy: a randomised, controlled, open-label, phase 3 trial. *Lancet*. 2017;390:849–860.
- Jacobson SG, Cideciyan AV, Ratnakaram R, et al. Gene therapy for leber congenital amaurosis caused by RPE65 mutations: safety and efficacy in 15 children and adults followed up to 3 years. *Arch Ophthalmol*. 2012;130:9–24.