

Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:

Maguire, MG;Birch, DG;Duncan, JL;Ayala, AR;Ayton, LN;Cheetham, JK;Cheng, P;Durham, TA;Ferris, FL;Hoyng, CB;Huckfeldt, RM;Jaffe, GJ;Kay, C;Lad, EM;Leroy, BP;Liang, W;McDaniel, LS;Melia, M;Michaelides, M;Pennesi, ME;Sahel, JA;Samarakoon, L

Title:

Endpoints and Design for Clinical Trials in USH2A-Related Retinal Degeneration: Results and Recommendations From the RUSH2A Natural History Study

Date:

2024-10-01

Citation:

Maguire, M. G., Birch, D. G., Duncan, J. L., Ayala, A. R., Ayton, L. N., Cheetham, J. K., Cheng, P., Durham, T. A., Ferris, F. L., Hoyng, C. B., Huckfeldt, R. M., Jaffe, G. J., Kay, C., Lad, E. M., Leroy, B. P., Liang, W., McDaniel, L. S., Melia, M., Michaelides, M., ... Samarakoon, L. (2024). Endpoints and Design for Clinical Trials in USH2A-Related Retinal Degeneration: Results and Recommendations From the RUSH2A Natural History Study. *Translational Vision Science and Technology*, 13 (10), <https://doi.org/10.1167/tvst.13.10.15>.

Persistent Link:

<https://hdl.handle.net/11343/359031>

License:

[CC BY-NC-ND](#)

Endpoints and Design for Clinical Trials in *USH2A*-Related Retinal Degeneration: Results and Recommendations From the RUSH2A Natural History Study

Maureen G. Maguire¹, David G. Birch², Jacque L. Duncan³, Allison R. Ayala¹, Lauren N. Ayton⁴, Janet K. Cheetham⁵, Peiyao Cheng¹, Todd A. Durham⁵, Frederick L. Ferris, III⁶, Carel B. Hoyng⁷, Rachel M. Huckfeldt⁸, Glenn J. Jaffe⁹, Christine Kay¹⁰, Eleonora M. Lad⁹, Bart P. Leroy¹¹, Wendi Liang¹, Lee S. McDaniel¹, Michele Melia¹, Michel Michaelides¹², Mark E. Pennesi^{2,13}, José-Alain Sahel^{14–16}, and Lassana Samarakoon¹, on behalf of the REDI Working Group and the Foundation Fighting Blindness Clinical Consortium Investigator Group

¹ Jaeb Center for Health Research, Tampa, FL, USA

² Retina Foundation of the Southwest, Dallas, TX, USA

³ University of California, San Francisco, San Francisco, CA, USA

⁴ University of Melbourne, and Centre for Eye Research Australia, East Melbourne, VIC, Australia

⁵ Foundation Fighting Blindness, Columbia, MD, USA

⁶ Ophthalmic Research Consultants, Waxhaw, NC, USA

⁷ Radboud University Medical Center, Nijmegen, Netherlands

⁸ Massachusetts Eye and Ear, Boston, MA, USA

⁹ Duke Department of Ophthalmology, Duke University, Durham, NC, USA

¹⁰ Vitreoretinal Associates, Gainesville, FL, USA

¹¹ Ghent University Hospital, Ghent University, Ghent, Belgium

¹² Moorfields Eye Hospital and UCL Institute of Ophthalmology, London, UK

¹³ Casey Eye Institute, Oregon Health & Science University, Portland, OR, USA

¹⁴ Institut de la Vision, Sorbonne Université, INSERM, CNRS, Paris, France

¹⁵ Centre Hospitalier National d'Ophthalmologie des Quinze-Vingts, Centre de Référence Maladies Rares REFERET and INSERM-DGOS CIC 1423, Paris, France

¹⁶ Department of Ophthalmology, The University of Pittsburgh School of Medicine, Pittsburgh, PA, USA

Correspondence: Allison Ayala, Jaeb Center for Health Research, 15310 Amberly Drive, Tampa, FL 33647, USA. e-mail:

ffbcorrespauth@jaeb.org

Received: April 2, 2024

Accepted: August 27, 2024

Published: October 9, 2024

Keywords: retinal degeneration; epidemiology; clinical trials

Citation: Maguire MG, Birch DG, Duncan JL, Ayala AR, Ayton LN, Cheetham JK, Cheng P, Durham TA, Ferris FL III, Hoyng CB, Huckfeldt RM, Jaffe GJ, Kay C, Lad EM, Leroy BP, Liang W, McDaniel LS, Melia M,

Purpose: To evaluate functional and structural assessments as endpoints for clinical trials for *USH2A*-related retinal degeneration.

Methods: People with biallelic disease-causing variants in *USH2A*, visual acuity $\geq 20/80$, and visual field $\geq 10^\circ$ diameter were enrolled in a 4-year, natural history study. Participants underwent static perimetry, microperimetry, visual acuity, fullfield stimulus testing (FST), and optical coherence tomography annually. Rates of change estimated from mixed-effects linear models and percentages of eyes with changes exceeding the coefficient of repeatability (CoR) or thresholds conforming with U.S. Food and Drug Administration (FDA) guidelines were evaluated.

Results: Rates of change were generally more sensitive to change than proportions of eyes exceeding a threshold such as the CoR. Baseline ellipsoid zone area $\geq 3 \text{ mm}^2$ was necessary to detect change. Mean sensitivity and volumetric hill of vision measures on static perimetry had similar properties and were the most sensitive to changes of the continuous measures. The highest 4-year proportions of eyes exceeding the CoR were from FST testing (47%) and microperimetry (32%). Specification of loci as functional transition points (FTPs) resulted in 45% (static perimetry) and 46% (microperimetry) at 4 years, meeting FDA guidelines for progression.

Michaelides M, Pennesi ME, Sahel JA, Samarakoon L. Endpoints and design for clinical trials in *USH2A*-related retinal degeneration: Results and recommendations from the RUSH2A natural history study. *Transl Vis Sci Technol.* 2024;13(10):15, <https://doi.org/10.1167/tvst.13.10.15>

Conclusions: Rate of change of mean sensitivity on static perimetry was a sensitive perimetric continuous measure. Percentages of within-eye change were largest with FST testing and microperimetry. FTPs appear to be particularly sensitive to change. These results affect clinical trial design for *USH2A*-related retinal degeneration.

Translational Relevance: Analyses of natural history data from the Rate of Progression in *USH2A*-Related Retinal Degeneration (RUSH2A) study can inform eligibility criteria and endpoints for clinical trials.

Introduction

The Foundation Fighting Blindness Clinical Consortium was established in 2016 as an international network of clinical centers to conduct research in participants with rare inherited retinal disorders (IRDs), with the goal to accelerate development of treatments for IRDs.^{1,2} The initial project for the Consortium was a natural history study of patients with biallelic disease-causing variants in the *USH2A* gene, one of the most common causes of IRDs. The resulting rod-cone degeneration may be accompanied by congenital hearing loss (Usher syndrome type 2 [USH2]) or not (non-syndromic autosomal recessive retinitis pigmentosa [ARRP]).³⁻⁵ Vision loss typically begins with nyctalopia followed by midperipheral visual field loss and slowly progresses toward both the center and periphery, often with a prolonged period of preservation of central vision.⁶⁻¹⁰

Two of the objectives of the Rate of Progression in *USH2A*-Related Retinal Degeneration (RUSH2A) study were to characterize the natural history of study eyes and to use the results to inform the design of clinical treatment trials intended to slow, stop, or reverse the progression of retinal degeneration. Previous publications from the RUSH2A study have addressed the first of these objectives by describing the findings from several imaging and functional tests at baseline and through 4 years.¹¹⁻¹⁷ The results presented in these publications and from additional analyses of the natural history data have been reviewed by the Consortium's Regulatory Endpoints and Trial Design for IRDs (REDI) Working Group. (Supplementary Appendix A). Members of the REDI Working Group have expertise in a broad range of topics, including genetics, clinical assessment of patients with IRDs, retinal imaging, multicenter clinical trials methodology, and biostatistics. The REDI Working Group also has shared data with and sought comments from representatives of the pharmaceutical industry involved in development of treatments for IRDs and the U.S. Food and Drug Administration (FDA) (Supplementary Appendix A). The purpose of this report is to

provide the results of a comprehensive evaluation by the REDI Working Group of the measures acquired during RUSH2A for suitability as outcome measures in clinical trials for *USH2A*-related retinal degeneration.

Methods

Design of the RUSH2A Study

The design and methods of the RUSH2A study have been described in detail in previous publications.¹¹⁻¹⁸ Participants were enrolled at 16 clinical sites in Europe and North America between August 2017 and December 2018. The institutional review boards or ethics committees associated with each participating site approved the consent process and study, which adhered to the tenets of the Declaration of Helsinki.

Study participants were 8 years of age or older and had a clinical presentation of rod-cone degeneration associated with at least two pathological variants in *USH2A*. The eye with the better best-corrected visual acuity (BCVA) was designated the study eye. Only participants having a study eye with an Early Treatment of Diabetic Retinopathy Study (ETDRS) letter score of 54 or greater (20/80 or better), central visual field at least 10° in diameter to a III4e target based on kinetic perimetry (KP), and stable fixation were enrolled in a primary cohort that was followed annually over 4 years.

Visual Function Testing and Imaging in the RUSH2A Study

Visual function testing and imaging were conducted by study-certified personnel following standard protocols at baseline and the annual visits through 4 years. Static perimetry (SP) was performed using the Octopus 900 automated perimeter (Haag-Streit, Koeniz, Switzerland) using the German Adaptive Thresholding Estimation (GATE) strategy with a size V stimulus and a custom, centrally weighted (CW), 185-point radial grid extending 65° nasally and superiorly, 67° inferiorly, and 80° temporally. In

addition to the mean sensitivity ($SP\ MS_{CW}$) in decibels (dB) across all points, the Casey Reading Center (Casey Eye Institute, Oregon Health & Sciences University, Portland, OR) generated a three-dimensional, quantitative surface map (hill of vision) from the SP values. The total volume beneath the surface (V_{TOT}), the volume below a central region with a radius of 30° (V_{30}), and the difference between V_{TOT} and V_{30} (V_{PERIPH}) was calculated in decibel-steradians (dB-sr).

Mesopic microperimetry (MP) was performed with the MAIA 2 Macular Integrity Assessment unit (CenterVue, Padova, Italy) using an 89-point grid arranged in concentric circles located at 2° , 4° , 6.5° , 9° , 12° , and 15° from the foveal center. Mean sensitivity in decibels (MP MS) was calculated. Fullfield stimulus testing (FST) was performed with the Espion E3 device (Diagnosys LLC, Lowell, MA) using white, blue, and red stimuli. Thresholds were reported in $\log\ cd\cdot s/m^2$.¹⁴ BCVA was measured after refraction with ETDRS charts or the electronic version of the test and summarized with a letter score from 0 (Snellen 20/800) to 100 (Snellen 20/10).¹⁹

Optical coherence tomography (OCT) was performed with the SPECTRALIS HRA+OCT (Heidelberg Engineering, Heidelberg, Germany). Acquisition parameters for each eye were as follows: one dense preset volume scan ($30^\circ \times 25^\circ$, 121 B-scans, automatic real-time tracking [ART set] to 9, high resolution) centered on the foveal center and one 7-line raster ($30^\circ \times 5^\circ$, 7 B-scans, ART set to 25, high resolution). Readers at the Duke Reading Center determined ellipsoid zone (EZ) area in square millimeters on OCT images.

Participants at all clinical sites had BCVA, SP, and OCT testing in the study eye. Testing of MP and FST in the study eye was conducted only at sites with the appropriate device. Central subfield thickness (CST) on OCT and responses on electroretinogram (ERG) testing were included in the original study protocol but are not included in this evaluation of possible endpoints for clinical trials. Macular edema changes CST dramatically and can be transient. If CST were an endpoint, interpretation of results would be difficult and possibly biased when some eyes developed macular edema. Responses on ERG were undetectable for a large proportion of patients at baseline, so that following ERG results would be non-informative with respect to detecting treatment group differences.¹⁴ Dark-adapted visual field testing and imaging using an adaptive optics scanning laser ophthalmoscope were conducted as part of ancillary studies at a small subset of clinical centers and are not included in the evaluation.^{16,18}

Data Analysis

Computations were performed using SAS 9.4 (SAS Institute, Cary, NC) or R (R Foundation for Statistical Computing, Vienna, Austria). Reported *P* values are two sided. The association between baseline values of each measure and duration of disease (time since onset of symptoms as reported by the participant) was assessed with Spearman correlation coefficients and are described in terms of negligible to very weak (0.00–0.19), weak (0.2–0.39), moderate (0.40–0.59), strong (0.60–0.79), and very strong (0.80–1.00) correlations as proposed by Evans.²⁰ Longitudinal linear regression models were used to estimate the mean annual rate of change of the outcome measures. Mixed-effects models with a random intercept were used to accommodate the correlation of observations over time within the same eye. For each outcome measure, only data from eyes that had at least two measurements, at least one of which was post-baseline, were included in the model. Time was calculated as the number of days from baseline divided by 365.25. A standardized annual change was calculated as the estimated slope over time (annual change) from the longitudinal model divided by both its standard error and the square root of the number of observations, with confidence intervals (CIs) calculated via bootstrapping. To protect the estimates from undue influence from outliers, additional models that down-weighted outlier individual rates of change were also applied. For these models, the rate of change for each participant was calculated from a simple linear regression model, then a robust regression model using M estimation with a Huber weighting function was used to calculate the weight to be applied in the mixed-effects model for each eye.^{21,22}

The percentage of eyes with a change from baseline exceeding the coefficient of repeatability (CoR) was calculated at 2 and 4 years after baseline.²³ Estimates of the CoR were derived from repeat testing during the baseline examination for SP, MP, and FST and from previous repeatability studies for visual acuity.^{11,13,19,24} Confidence intervals for proportions were calculated using the Wilson score method.

In addition, the FDA has provided guidance on clinically meaningful changes within an eye for SP and MP (mean change of ≥ 7 dB in ≥ 5 prespecified points) and BCVA (change in letter score of ≥ 15).^{25–28} The percentages of eyes with changes of these magnitudes were calculated at 2 and 4 years. For the perimetric measures, the value of a point at baseline had to be ≥ 8 dB to be considered a candidate for a prespecified point. One approach for specifying the points on MP and SP was based on results by Hood et al.²⁹ compar-

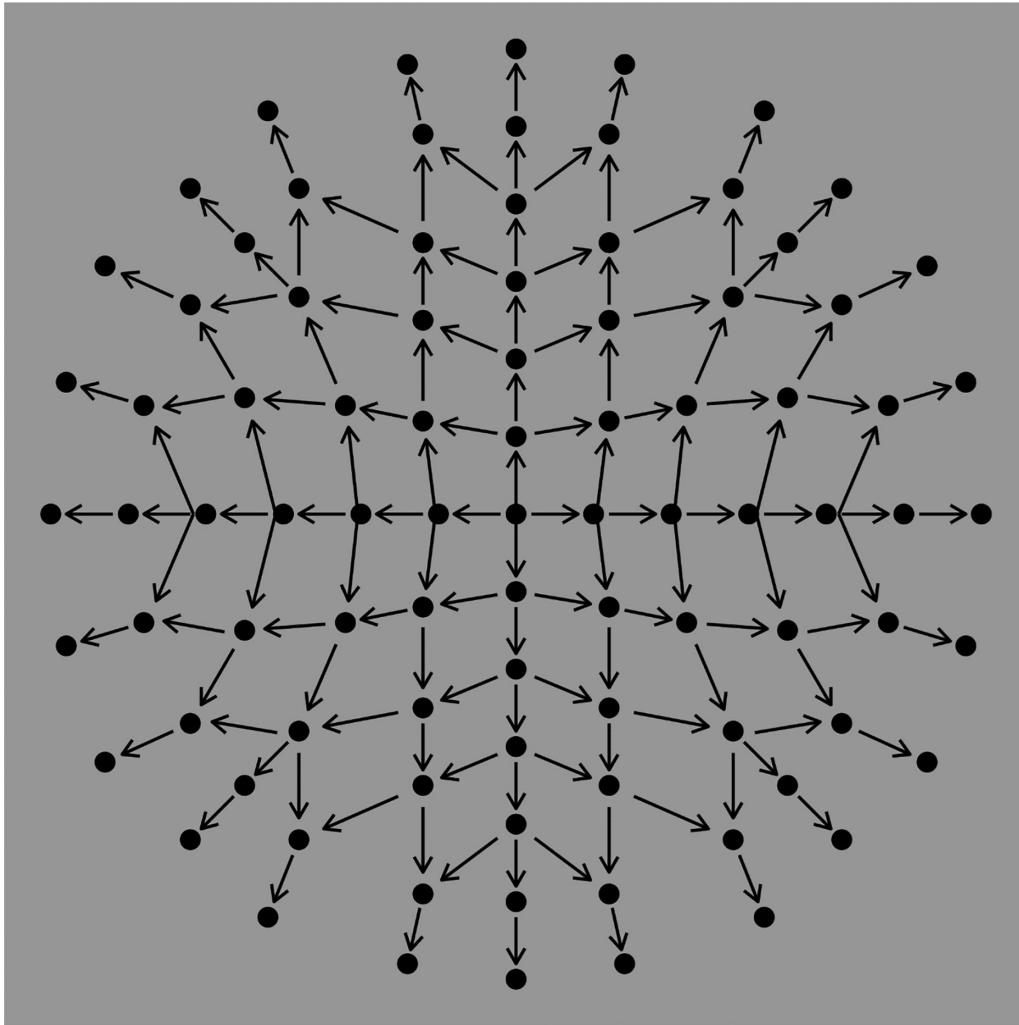


Figure 1. Pathways on the microperimetry testing grid for selecting FTPs. The schematic of the testing grid for microperimetry shows pathways from each point to more peripheral adjacent points. FTPs were identified by comparing each point to the 1 to 4 more peripheral adjacent points on the testing grid. A decrease in sensitivity of ≥ 7 dB from an inner point to the more peripheral adjacent point qualified the inner point as a candidate FTP. Candidate FTPs were ordered by the percentage of qualifying adjacent points, from 100% to 25%. Candidate points that were qualified by 100% of adjacent points were selected as FTPs. If the selected number of selected points was less than 5, then points with the next highest percentage of qualifying points were included.

ing the sensitivity of points on perimetric testing to the retinal location of the EZ on OCT. They found that a large decrease in sensitivity between adjacent points along a pathway from the center to periphery indicated a likely transition from within the EZ to outside the EZ. Functional transition points (FTPs) for static perimetry and microperimetry were identified by comparing each point to one to four (depending on the location of the point in the testing grid) more peripheral adjacent points on the testing grid (Fig. 1). When there was a decrease in sensitivity of ≥ 7 dB from an inner point to the more peripheral adjacent point, the inner point was qualified as a candidate FTP. All candidate FTPs were ordered by the percentage of qualifying adjacent points, from 100% to 25%. Candidate points that were

qualified by 100% of adjacent points were selected as FTPs. If the selected number of selected points was less than 5, then points with next highest percentage of qualifying points were included. Other approaches for prespecification of points were evaluation of the entire set of points on the perimetric testing grid and of the points in the central 30° (SP only).

Upon inspection of the data, we observed that some eyes with baseline values indicating severe degeneration on SP, MP, or OCT did not change over time, either because no further deterioration occurred, or the changes were so small that they could not be detected.¹⁷ Because enrollment of such eyes would be unlikely in a clinical trial of a treatment to reduce further progression as assessed by the specific measure,

Table 1. Desirable Properties of Outcome Measures in Clinical Trials for Inherited Retinal Disorders**Assessment procedure**

- Accurate—The procedure measures the intended characteristic
- Reproducible—Results are similar with repeat testing when no change occurs in the patient condition
- No floor or ceiling effects—Measurement range covers the study population
- Objective—Results are not under the control of the patient or examiner
- No learning or fatigue effects

Association with the retinal degeneration

- Correlates with disease severity or disease duration
- Sensitive to change in disease severity or duration
 - Continuous measures—Mean change over time is large compared to the variability of change across study participants (high signal-to-noise ratio)
 - Discrete events within an eye—The change is greater than measurement error
- Clinically relevant—Measures a characteristic that matters to patients
- If not directly clinically relevant, changes are closely correlated to changes in clinically relevant measures (requirement for surrogate outcome)

Feasibility

- Available at all clinical sites
- Can be administered by the study personnel
- Can be performed by the study population (e.g., very young, very old)
- Relatively low burden in terms of time, effort, and difficulty to patients
- Low cost

preserved cohorts were defined by determining a threshold baseline value above which the slopes of the measure over time was greater than 0: $V_{TOT} > 5$ dB-sr, $V_{30} > 3$ dB-sr, $V_{PERIPH} > 5$ dB-sr, $SP\ MS_{CW} > 4$ dB, $MP\ MS > 1$ dB, and $EZ\ area > 3\ mm^2$.

Properties of Outcome Measures

The REDI Working Group developed a set of desirable properties for outcome measures for IRDs related to the assessment procedure, association with the retinal degeneration, and feasibility in a clinical trial setting (Table 1). Properties associated with *USH2A*-related retinal degeneration and the range of measurement are the main focus of this evaluation. Other properties of the measures were considered in the evaluation only when they were particularly favorable or particularly unfavorable.

Because of the limited population of people affected with *USH2A*-related retinal degeneration, sensitivity to change in disease severity or duration was considered particularly important because of its strong influence on required sample sizes for clinical trials. At a given degree of treatment efficacy, continuous measures that are more sensitive to change in disease severity (steeper

slope over time) change to a similar degree across individual people regardless of baseline values (similar slope for each person), and they have little measurement error (relatively low CoR), thus yielding lower sample sizes for clinical trials. The standardized rate of change captures these components and has no units, allowing for comparison across measures when they are captured with the same examination frequency.¹⁷ Measures that require exceeding a within-eye specified change threshold must yield a percentage in untreated eyes that allows room to demonstrate a treatment effect; the treatment effect could be a decrease in the percentage of eyes that worsen or an increase in the percentage of eyes that remain stable or improve. Measures derived from test procedures that do not place a heavy time or effort burden on the patient are more desirable than those involving a heavy burden; descriptors of relative patient burden were based on the opinions of the members of the REDI Working Group.

Sample Size Calculations

Sample sizes for hypothetical scenarios for clinical trials comparing rates of change between treatment groups were calculated assuming two groups with equal

allocation between treatment and control with a target of 80% power ($\beta = 0.2$) and two-sided type 1 error probability of 5% ($\alpha = 0.05$). The null hypothesis being tested was the hypothesis that the coefficient associated with the interaction between time and treatment was equal to zero (i.e., no difference between treatment groups) in a random intercept model with fixed effects of time and the interaction between time and a treatment indicator. The alternative hypothesis was specified as a multiple (c) of the standardized rate of change (γ), and the total sample size (n) was calculated as

$$n = \frac{(Z_{1-\alpha/2} + Z_{1-\beta})^2}{\gamma^2 c^2 \pi (1 - \pi)}$$

where $Z_{1-\alpha/2}$ is the $1 - \alpha/2$ quantile of the standard normal distribution, $Z_{1-\beta}$ is the $1 - \beta$ quantile of the standard normal distribution, c is a constant multiple, π is the proportion of subjects assigned to the treatment group, and β is the required type 2 error probability.

Results

Study Population

Among the 105 participants in the primary cohort, two died during the follow-up period. Among the 103 participants with at least two visits to a clinical center, the clinical diagnosis was USH2 for 64 participants (62%) and ARRP for 39 participants (38%). The mean \pm SD age was 37 ± 12 years, and 58 of the participants were female (56%). Median duration of disease as reported by the participant at enrollment was 12 years (interquartile range [IQR], 7–20). Demographic characteristics were similar in the MP cohort ($n = 94$) and the FST cohort ($n = 77$). The visit completion rates were 102/104 at 1 year (98%), 88/104 at 2 years (85%), 99/104 at 3 years (95%), and 95/103 at 4 years (92%). Many of the missed visits at 2 years were attributable to the suspension of clinical research studies during the COVID-19 pandemic. The number of eyes used in analyses varied depending on the testing modality because testing was conducted only in centers that had the required equipment for MP and FST and because of ungradable OCT images (Fig. 2). Preserved cohorts for each measure excluded eyes with no or little change over time.

Measures From Static Perimetry

At baseline, the SP summary measures of V_{TOT} , V_{30} , V_{PERIPH} , and SP MS_{CW} were very strongly correlated (Spearman's rank correlation coefficient, $r_s \geq 0.82$)

with each other (Supplementary Table S1), and each was similarly correlated with duration of disease, with r_s values ranging from -0.33 to -0.43 (Supplementary Table S2). In the preserved cohorts, the 4-year changes in the measures were highly correlated between V_{TOT} and V_{PERIPH} ($r_s = 0.96$) and between V_{30} and SP MS_{CW} ($r_s = 0.90$). The percentage of the entire cohort that qualified for the preserved cohort was similar (approximately 90%) for V_{TOT} , V_{30} , and SP MS_{CW} and lower (77%) for V_{PERIPH} (Table 2). The standardized rate of change was similar for the four measures, ranging from -1.28 for V_{PERIPH} to -1.58 for SP MS_{CW} . The comparison of features for the entire cohort is given in Supplementary Table S2. When only the points in the FTPs were considered, the standardized rate of change was -1.94 (95% CI, -2.09 to -1.79) for SP MS_{CW} .

The percentages of eyes with worsening greater than the CoR were similar among the four measures at 2 years (8%–12%) and 4 years (19%–27%) (Table 2). The percentages of eyes with improvement greater than the CoR was 1% at 2 years for all four measures and 0% at 4 years for all measures except V_{PERIPH} (1%). The percentages of eyes that met the FDA guidelines for clinically meaningful worsening for SP when all points on the grid were evaluated for prespecification were 2% at 2 years and 5% at 4 years; the corresponding percentages were 0% and 9%, respectively, when only the points within the 30° field were evaluated for prespecification (Table 3). When the FTPs selected from all the points on the grid were considered as the prespecified points, the percentage worsening was 31% at 2 years and was 45% at 4 years. Reliability associated with the FTPs was assessed by identifying FTPs based on the results from the first testing session at baseline and comparing them to the results from the second testing session at baseline. The percentage of eyes classified as worsening on the second testing session was 10%, and the CoR for these FTPs was 16.9 dB. For all three approaches to prespecifying points, the percentage of eyes with improvement at 2 or 4 years was 0% or 1%.

Measures From Microperimetry

The correlation (r_s) of the MP MS with duration of disease at baseline was -0.33 (Supplemental Table S2). The percentage of the entire cohort that qualified for the preserved cohort was 91%. The standardized rate of change was -1.01 . When only FTPs were considered, the standardized rate of change was -1.79 (95% CI, -1.97 to -1.61). The percentages of eyes with worsening greater than the CoR were 14% at 2 years and 33% at 4 years. The percentages of eyes with improvement greater than the CoR were 1% at 2 years and 3% at

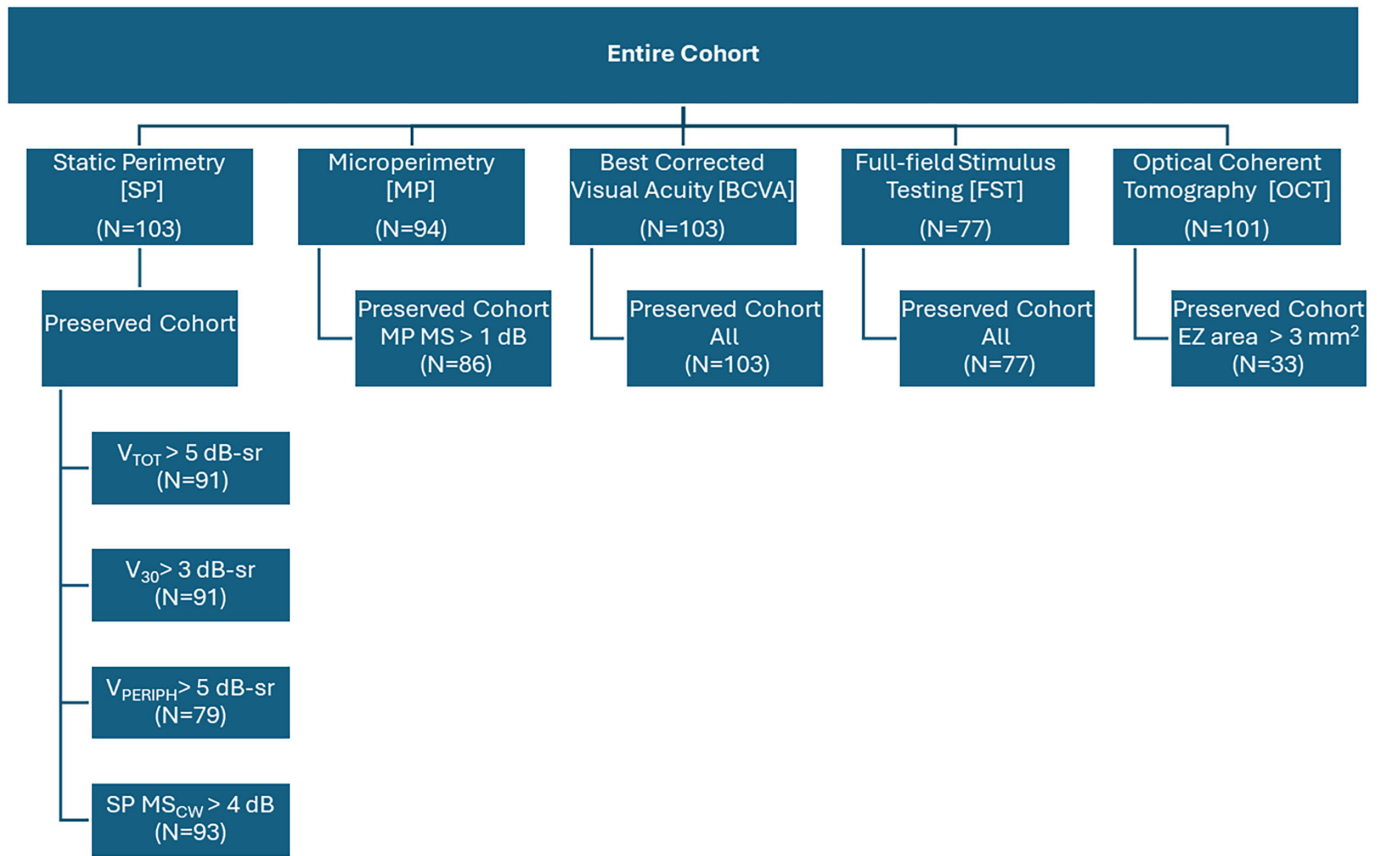


Figure 2. Participants included in analyses for each measure. Shown are the number of participants in the entire cohort and the preserved cohort for each measure and testing modality. V_{TOT} is the total volume beneath the surface of the hill of vision, V_{30} is the volume below a central region with radius of 30°, and V_{PERIPH} is $V_{TOT} - V_{30}$.

translational vision science & technology

4 years. The percentages of eyes that met the FDA guidelines for clinically meaningful worsening when all points on the MP grid were considered as prespecified were 0% at 2 years and 5% at 4 years (Table 3). When the FTPs were considered as the prespecified points, the percentage worsening was 19% at 2 years and was 46% at 4 years, whereas no eyes exhibited clinically meaningful improvement at 2 or 4 years. Reliability associated with the FTPs was assessed by identifying FTPs based on the results from the first testing session at baseline and comparing them to the results from the second testing session at baseline. The percentage of eyes classified as worsening on the second testing session was 7% and the CoR for these FTPs was 11.3 dB.

Best-Corrected Visual Acuity

The correlation (r_s) of BCVA score with duration of disease at baseline was -0.26 (Table 2). There was no floor effect observed in the data, so 100% of eyes qualified for the BCVA preserved cohort. The annual rate of change was -0.80 letters, and the standardized

rate of change was -0.89 . The percentages of eyes with worsening greater than the CoR was 5% at 2 years and 12% at 4 years. The percentages of eyes with improvement greater than the CoR were 1% at 2 years and 0% at 4 years. The percentages of eyes that met the FDA guidelines for clinically meaningful worsening for BCVA were 0% at 2 years and 2% at 4 years (Table 3). No eyes exhibited clinically meaningful improvement at 2 or 4 years.

Fullfield Stimulus Testing

The thresholds from testing with white and blue stimuli were very strongly correlated at baseline ($r_s = 0.97$), and the changes in threshold from baseline at 4 years between white and blue stimuli were highly correlated ($r_s = 0.74$). This is consistent with both being detected by the same receptor type (rods in eyes that retained them). Results from testing with red stimuli are mediated by cones in a majority of patients and show a slower rate of decline than for the white and blue stimuli. Therefore, only the FST threshold results

Table 2. Features of Outcome Measures Using Values From the Preserved Cohort for Each Measure

Features of Outcome Measures	Outcome Measures									
	V_{TOT}	V_{30}	V_{PERIPH}	SP MS _{CW}	MP MS	BCVA	FST White	EZ Area		
Unit of measurement	dB-sr	dB-sr	dB-sr	dB	dB	Letters	log cd·s/m ²	Mm ²		
Baseline correlation with duration (r_s) ^a (95% CI)	-0.24 (-0.43 to -0.03)	-0.27 (-0.45 to -0.07)	-0.15 (-0.36 to 0.08)	-0.29 (-0.46 to -0.09)	-0.31 (-0.49 to -0.11)	-0.26 (-0.44 to -0.07)	0.51 (0.32 to 0.66)	-0.24 (-0.54 to 0.12)		
Total cohort in preserved cohort (%)	88	88	77	90	91	100	100	33		
Eyes in longitudinal model, n	91	91	79	93	86	103	77	33		
Annual rate of change ^b (95% CI)	-2.02 (-2.33 to -1.71)	-0.54 (-0.62 to -0.47)	-1.71 (-2.00 to -1.41)	-0.54 (-0.61 to -0.47)	-0.35 (-0.42 to -0.27)	-0.80 (-0.97 to -0.62)	0.09 (0.07 to 0.11)	-0.34 (-0.47 to -0.22)		
Standardized rate of change ^b (95% CI)	-1.34 (-1.66 to -1.08)	-1.54 (-1.82 to -1.31)	-1.28 (-1.68 to -0.99)	-1.58 (-1.86 to -1.35)	-1.01 (-1.33 to -0.71)	-0.89 (-1.13 to -0.66)	1.09 (0.79 to 1.44)	-0.95 (-1.49 to -0.49)		
CoR	14.0	3.4	11.4	3.5	2.2	8.8	0.39	N/A		
Eyes with data at 2 y, n	77	78	68	78	70	88	69	29		
Eyes worsening \geq CoR at 2 y (%) (95% CI)	8 (4 to 16)	8 (4 to 16)	9 (4 to 18)	12 (6 to 21)	14 (8 to 24)	5 (2 to 11)	30 (21 to 42)	N/A		
Eyes improving \geq CoR at 2 y (%)	1	1	1	1	1	1	3	N/A		
Eyes with data at 4 y, n	78	80	67	79	72	95	66	31		
Eyes worsening \geq CoR at 4 y (%) (95% CI)	21 (13 to 31)	24 (16 to 34)	27 (18 to 39)	19 (12 to 29)	33 (24 to 45)	12 (7 to 20)	47 (35 to 59)	N/A		
Eyes improving \geq CoR at 4 y (%)	0	0	1	0	3	0	3	N/A		
Direct clinical relevance	High	High	High/medium	High	High	High	Low	High		
Subjectivity (patient or examiner)	High	High	High	High	Medium	Medium	Medium	Low		
Relative patient burden	High	High	High	High	High	Medium	Medium	Low		

N/A, not available.

^aDuration is time since onset of symptoms as reported by the participant.

^bBased on the 4-year longitudinal model with outliers down-weighted for each measure.

Table 3. Eyes Meeting FDA Guidelines for Clinically Meaningful Changes in SP, MP, and BCVA

Outcome Measure	2 Years					4 Years				
	n	Worsening		Improvement		n	Worsening		Improvement	
		%	95% CI	%	95% CI		%	95% CI	%	95% CI
Static perimetry^a										
Entire grid	85	2	0.7–8	0	0–4	91	5	2–12	1	0.03–6
Central 30° (108 points)	85	0	0–4	0	0–4	91	9	5–16	1	0.03–6
Functional transition points ^b	85	31	22–41	0	0–4	91	45	35–55	0	0–4
Microperimetry^a										
Entire grid	77	0	0–5	0	0–5	75	5	2–13	0	0–5
Functional transition points ^c	70	19	11–29	0	0–5	70	46	35–57	0	0–5
BCVA	88	0	0–4	0	0–4	95	2	0–7	0	0–4

^aOnly points with a baseline value of ≥ 8 dB are included and only eyes with ≥ 5 points at baseline meeting the region definition are included.

^bThe number of functional transition points ranged from 5 to 30 for eyes analyzed at 2 years and from 5 to 30 for eyes analyzed at 4 years.

^cThe number of functional transition points ranged from 5 to 22 for eyes analyzed at 2 years and from 5 to 18 for eyes analyzed at 4 years.

Table 4. Sample Size and Power Estimates for Clinical Trials of Treatments Intended to Slow Progression

Untreated Group	Example Measures ^a	Treated Group Relative Reduction, Total Sample Size (n) ^b		
		30%	60%	100%
Standardized rate of worsening				
2 y of follow-up				
0.40	BCVA, EZ	2181	546	197
0.70	MP MS, FST	712	178	65
0.80	V _{TOT} , SP MS _{CW}	546	137	50
4 y of follow-up				
1.00	MP MS, FST, EZ	349	88	32
1.25	V _{TOT}	224	56	21
1.50	SP MS _{CW}	156	39	14
1.75	SP FTP, MP FTP	114	29	11
Eyes worsening more than CoR or FDA threshold				
2 y of follow-up				
5%	BCVA ^c	5676	1176	304
10%	SP MS _{CW} ^c	2711	566	147
15%	MP MS, ^c MP FTP ^d	1723	362	95
30%	FST, ^c SP FTP ^d	734	159	43
4 y of follow-up				
20%	V _{TOT} ^c	1229	261	69
25%	V ₃₀ ^c	932	200	53
30%	MP MS ^c	734	159	43
45%	FST, ^c SP FTP, ^d MP FTP ^d	405	91	25

^aStandardized rate of worsening in the preserved cohorts.

^bAdditional assumptions: 80% power; one active treatment group and one untreated group; equal (1:1) allocation between groups; annual examinations; and no loss to follow-up.

^cPercentage worsening by the CoR or more.

^dPercentage worsening by the threshold recommended in FDA guidance for SP and MP points prespecified as the FTPs.

Table 5. Sample Size and Power Estimates for Clinical Trials of Treatments Intended to Reverse Progression

Untreated Group	Example Measures ^a	Treated Group, Relative Rate of Improving, ^a Total Sample Size (n) ^b		
		10%	20%	30%
Standardized rate of worsening				
2 y of follow-up				
0.40	BCVA, EZ	163	137	117
0.70	MP MS, FST	53	45	38
0.80	V _{TOT} , SP MS _{CW}	41	35	30
4 y of follow-up				
1.00	MP MS, FST, EZ	26	22	19
1.25	V _{TOT}	17	14	12
1.50	SP MS _{CW}	12	10	9
1.75	SP FTP, MP FTP	9	8	7
		Treated Group, Percentage Improving (n)		
Untreated Group Percentage Improving		20%	30%	50%
0%		69	43	22
5%		151	71	29
10%		398	124	39

^aExample: If the standardized rate of worsening in the untreated group was 0.40 and the relative rate of improvement was 10%, then the standardized rate of improving in the treated group is 0.04.

^bAdditional assumptions: 80% power; one active treatment group and one untreated group; equal (1:1) allocation between groups; annual examinations; and no loss to follow-up.

from testing with white stimuli were included in this analysis.

The correlation (r_s) of the FST threshold with duration of disease at baseline was 0.51 (Table 2). There was no floor effect observed in the data, so 100% of eyes qualified for the FST preserved cohort. The standardized annual rate of change was 1.09. The percentages of eyes with worsening greater than the CoR were 30% at 2 years and 47% at 4 years. The percentage of eyes with improvement greater than the CoR was 3% at both 2 years and 4 years.

Ellipsoid Zone Area

The correlation (r_s) of the EZ area threshold with duration of disease at baseline was -0.50 (Supplementary Table S2). The percentage of the entire cohort that qualified for the preserved cohort was 33%. In the preserved EZ cohort, the annual rate of change was -0.34 mm^2 , and the standardized rate of change was -0.95 .

Features of Measures for the Entire Cohort

The comparison of features for the entire cohort, including eyes that exhibited no change during follow-

up, is given in Supplementary Table S2. Overall, the annual rates of change, the standardized rates of change, and the percentages worsening more than the CoR at 2 and 4 years were lower in the entire cohort than in the preserved cohort, except when 100% of eyes qualified for the preserved cohort. Most associations between measures remained similar in the entire cohort.

Sample Size Requirements for Clinical Trials

The results of sample size calculations for clinical trials designed to detect reductions in the rate of worsening are displayed in Table 4. Examples of measures with rates for untreated eyes that were similar to the table values are provided. For clinical trials aimed at reducing the rate of progression within 2 years of follow-up, the required sample size ranges were 546 to 2181 for a 30% reduction, 137 to 546 for a 60% reduction, and 50 to 197 for a 100% reduction (stopping progression). For 4 years of follow-up, the required sample size ranges were 114 to 349 for a 30% reduction, 29 to 88 for a 60% reduction, and 11 to 32 for a 100% reduction (stopping progression). The required sample size ranges for a 2-year trial aimed at reduc-

ing the percentage of eyes with worsening exceeding a threshold value such as the CoR were 734 to 5676 for a 30% reduction, 159 to 1176 for a 60% reduction, and 43 to 304 for a 100% reduction. The required sample size ranges for a 4-year trial aimed at reducing the percentage of eyes with worsening exceeding a threshold value were 405 to 1229 for a 30% reduction, 91 to 261 for a 60% reduction, and 25 to 69 for a 100% reduction.

The results of sample size calculations for clinical trials designed to detect improvement are displayed in Table 5. For 2 years of follow-up, the required sample size range was 41 to 163 for a 10% relative rate of improvement, 35 to 137 for a 20% relative rate of improvement, and 30 to 117 for a 30% relative rate of improvement. The required sample size ranges for comparing the percentage improving by more than a threshold value, such as the CoR, in treated eyes compared to untreated eyes was 69 to 398 for a 10% improvement in treated eyes, 43 to 124 for a 20% improvement, and 22 to 39 for a 30% improvement when the percentage improving in the untreated group ranged from 0% to 10%.

Discussion

The RUSH2A study has provided a wealth of information on the natural history of *USH2A*-associated retinal degeneration using state-of-the-art methods to assess retinal function and structure. Strengths and weaknesses of the studied measures for suitability as outcome measures in clinical trials varied and depended on the (1) intended effect of the candidate intervention (slow, stop, or reverse progression) and (2) whether the difference in average change between groups of eyes or the difference in proportions of individual eyes meeting a threshold amount of change would be used for the comparison of treatment groups.

Evaluation of Outcome Measures for Clinical Trials of Interventions Intended to Slow or Stop Progression

One requirement for eyes enrolled in clinical trials of interventions intended to slow or stop progression is that they have the capacity to demonstrate progression, or worsening, of the outcome variable when untreated during the trial follow-up period. Although the RUSH2A eyes chosen for longitudinal assessment had a central visual field at least 10° in diameter to a III4e target based on kinetic perimetry, worsening over 4 years was not detected for all eyes for all

measures. Only eyes that could worsen from baseline were included in the preserved cohort for each measure. Worsening was detectable in BCVA and FST white for all eyes, regardless of the baseline level. Approximately 10% of eyes did not worsen for each of the perimetric measures; however, in the 67% of eyes with small baseline EZ area (<3 mm²), EZ area did not decrease. EZ area has highly desirable features for the following reasons: It is an objective measure, it has high clinical relevance because it indicates the extent of fairly normal retinal sensitivity, and it has relatively high correlation with duration of symptoms. However, use of the EZ area as a primary outcome measure restricts the study population to those with an area at least 3 mm², thereby reducing the generalizability of results and increasing the difficulty of patient recruitment. In addition, among the 33% of eyes with a baseline EZ area >3 mm², the standardized rate of change was lower than the rates for all other measures except for BCVA.

Progressive loss of rod function and constriction of the visual field with relative preservation of central retinal structure and function until late in the disease process are hallmarks of *USH2A*-associated retinal degeneration (and all rod-cone dystrophies). Thus, perimetric measures were expected to characterize well the progression of degeneration. The mean sensitivity (SP MS_{CW}) and three volumetric measures (V_{TOT}, V₃₀, V_{PERIPH}) derived from the centrally weighted static perimetry testing grid were similar with respect to correlation with duration of disease, standardized rate of change, and percentages exceeding the CoR. Given the high degree of correlation among these measures at baseline and among changes from baseline to 4 years, such similar performance is not surprising. However, SP MS_{CW} requires only averaging the threshold values from the 185 points on the grid, whereas the three volumetric measures require implementation of more complicated, costly, and time-consuming algorithms. In the absence of an advantage to the volumetric measures, SP MS_{CW}, which incorporates threshold values throughout an approximate 140° visual field, is preferred as an outcome measure from among the four SP summary measures. In addition, SP MS_{CW} had one of the highest standardized rates of change among all the measures that were studied.

Microperimetry testing of the central 30° visual field allows accurate and precise placement of the stimulus on the retina, as reflected in the relatively low CoR (2.2 dB). Although MP MS did not have a higher standardized rate of change than SP MS_{CW}, the higher precision in placement of the stimuli in MP testing during follow-up than in SP testing contributed to somewhat higher percentages of eyes worsening more

than the CoR threshold at 2 and 4 years than for the SP summary measures.

As noted above, preservation of central visual function despite substantial loss of peripheral visual function was expected, and the RUSH2A data conformed to this expectation. Although the RUSH2A eligibility criteria for longitudinal follow-up specified an ETDRS letter score of 54 letters or greater (20/80 or better) for BCVA, the worst baseline BCVA score was 64 letters (20/50) and the mean was 81 letters (20/25). The annual rate of change was less than 1 letter per year (0.8 letter), yielding the lowest estimates of standardized rate of change and the lowest percentages of eyes with decrease exceeding the CoR at 2 and 4 years. Thus, although visual acuity is an established, highly clinically relevant, widely available outcome measure that does not place a high burden on the patient, the slow rate of decline leaves little possibility for a new treatment to demonstrate slowing or stopping the rate of decline or reducing the percentage of eyes meeting an accepted threshold for worsening within 4 years or less.

FST testing was developed specifically to assess visual function in eyes with severe vision loss without stable fixation.^{29,30} The thresholds from FST testing are dominated by the most sensitive retinal region under dark-adapted conditions, but, as a fullfield test, they do not enable retinal localization. We found that the white stimulus FST thresholds worsened over time not only for eyes with poor baseline results but also throughout the wide range of baseline values in the study eyes (4.7 log units). The low CoR in our study (0.39 log cd·s/m²), consistent with estimates of the CoR in other studies,^{30–32} is a particular strength for measures of the proportion of eyes that change by a threshold amount over time. At 2 years, 30% of eyes had worsened beyond the CoR and 47% by 4 years. The low percentage of eyes (3%) with improvement beyond the CoR demonstrates that the high percentage with worsening is not a manifestation of high random variability among measurement sessions. Although FST does not offer an advantage over the perimetric measures in the mean rate of worsening in a group of eyes, it is one of the most sensitive measures for detecting worsening visual function over time in a given eye.

In previous studies, the mean change in white stimulus FST threshold corresponded with changes in performance on the multi-luminance mobility test, the primary outcome measure in the phase 3 clinical trial of voretigene neparvovec for treatment of retinal degeneration caused by *RPE65* mutations.³³ The European Medicines Agency viewed the improvement in FST thresholds as supportive toward the efficacy of the

treatment.³⁴ However, the impact of changes in FST on functional vision has not been established for patients with good central vision.

Percentages of Eyes Meeting Criteria for Clinically Meaningful Worsening, According to FDA Guidance

When the criteria for clinically meaningful change, as recommended by the FDA, were applied to the RUSH2A results, very small percentages of eyes met the criteria for worsening at 2 years ($\leq 2\%$) or 4 years ($\leq 9\%$) for BCVA and the perimetric measures involving the entire visual field or the central visual field (Table 3). In this cohort, the FDA criteria are not sensitive to the changes that occur within 4 years or less in these measures.

Two perimetric measures provided some opportunity for a treatment to demonstrate a reduction in worsening as recommended by FDA guidance. Both mean change in MP sensitivity and in SP mean sensitivity when only the FTPs were prespecified and included in the calculation of worsening yielded percentages greater than 40% at 4 years. Nonetheless, the required sample sizes for a treatment that provides a 60% reduction in the percentage of eyes worsening by FDA criteria are at least 150 for a 2-year trial and 90 for a 4-year trial. In addition, because of the lower reliability of threshold results at FTPs, there may be a small percentage ($\sim 10\%$) of eyes classified as worsening when no change has occurred. Also, for localized treatments such as subretinal gene therapy, the approach to identifying the points most likely to change for FTP analyses would have to be modified to account for the changes anticipated in the area of direct treatment and surrounding areas.

Evaluation of Outcome Measures for Clinical Trials of Interventions Intended to Reverse Progression

All measures evaluated had an estimated rate of change indicating progression and none had a percentage of eyes greater than 5% with improvement that exceeded the CoR. Thus, improvements on either the group or eye level among treated eyes are unlikely to be induced by measurement error or true small fluctuations in function. Only small rates of improvement in a treated group are needed when comparing to a rate of worsening in an untreated group. For example, with 2 years of follow-up, if the rate of change for SP MS_{CW} is assumed to be -0.55 dB per year in the untreated group and $+0.11$ dB per year in the treated

group, the total sample size required is approximately 35. If the percentage improving more than the CoR is assumed to be 5% in the untreated group and 20% in the treated group, the total sample size required is approximately 71.

Use of Rate of Change for the Group Versus Percentage of Eyes With Change Exceeding the CoR

Examination of Table 4 shows that sample sizes are substantially larger (often 2 times or more) when the proportion exceeding the CoR is used as the outcome measure for treatment group comparison rather than the rate of change. These findings are similar to those found when assessing sample size considerations for clinical trials for treatment of glaucoma.^{35,36} Generally, dichotomizing an outcome causes a loss of statistical efficiency.³⁷ In addition, the longitudinal model used to estimate the rate of change uses not only the baseline and final measurement but also all intermediate measurements, further increasing statistical efficiency. The possible exception to the advantage of using the rate of change is FST, where the high repeatability of the test contributes to the efficiency of the comparison of proportions.

Reducing the rate of progression has important clinical implications. Our results and those of previous investigators of *USH2A*-related degenerations show that the loss of visual function is slow but unrelenting.⁶⁻¹⁰ Although some patients report a sudden decrease in their vision associated with the onset of inability to perform specific activities, steady visual function over time followed by a precipitous decrease was rarely observed in the RUSH2A participants. This discordance between patient report and the measurements of visual function may be attributed to patients maintaining a functional reserve of vision (i.e., ability greater than the level needed to perform an activity) that allows them to perform the activity. When the decrease in functional reserve crosses the threshold for being able to perform the activity, the patient becomes acutely aware of the loss in function.³⁸

Limitations and Caveats

We used the CoR as the cutpoint for dichotomizing changes from baseline. Although this cutpoint performed well with respect to limiting classification of eyes as improving, other cutpoints may perform equally well or better. In fact, because the percentage of eyes with improvement was so low at 2 years, cutpoints requiring less worsening should be explored to increase the percentage of eyes classified as worsen-

ing. However, worsening in the untreated group in a clinical trial setting may be less than in a natural history study because clinical trial participants may be masked to treatment and subject to the placebo effect. Similarly, improvement may be greater in untreated groups in a clinical trial. The sample size and power calculations presented in this paper allow comparison among outcome measures and among the magnitude of treatment effects for a specific set of assumptions regarding study population, number of treatment groups, frequency of examinations, and statistical power. Calculations for a specific clinical trial should be customized to the specific trial design.

Summary Recommendations

Evaluation of treatments for *USH2A*-related retinal degeneration (and the vast majority of other IRDs) faces the challenges of a relatively limited population of patients, slow rates of progression, and high variability in the rates of progression across patients, even within a group with disease-causing variants in the same gene. Use in clinical trials of outcome measures that are sensitive to change and clinically important is required to identify beneficial interventions.

Based on the data from the RUSH2A study, the REDI Working Group has the following recommendations:

- Rates of change (slopes) are more sensitive to change than the proportions of eyes exceeding a threshold such as the CoR. They should be the top choice for primary efficacy outcome measures. However, use of slopes may not be appropriate if both responders and non-responders are expected or if serious complications (e.g., retinal detachment) may occur that would markedly affect the slope in the treated group, masking a beneficial effect for most patients.
- Use of the rate of change in SP MS_{CW} as estimated through longitudinal regression methods provides an outcome measure that is sensitive to change. The volumetric summary measures V_{TOT} and V₃₀ have similar properties but offer no advantage over SP MS_{CW}.
- Eyes enrolled in clinical trials for treatment of *USH2A* should have an EZ area large enough (at least 3 mm²) to detect changes in the EZ area over time.
- MP is better than SP in detecting the proportion of eyes changing by the CoR. However, MP and SP detected similar percentages with change accord-

ing to criteria recommended by the FDA when considering only FTPs after 4 years.

- FST test results are sensitive to change over time, particularly for the proportion of eyes changing by more than the CoR. Although results of FST testing were cited as supporting evidence for the approval of voretigene neparvec by the European Medicines Agency for patients with *RPE65*-mediated inherited retinal dystrophy,³³ further studies are needed to understand the clinical impact of changes in FST in patients with good central visual function. However, FST may now be useful in detecting the effects of treatment in early-phase clinical trials.

Acknowledgments

Supported by the Foundation Fighting Blindness; by National Eye Institute, National Institutes of Health, Core Grants for Vision Research (EY002162 to JLD; EY010572 to MEP; EY08098 to J-AS); by unrestricted grants from Research to Prevent Blindness (JLD, MEP, J-AS); by an Australian National Health and Medical Research Council Investigator Grant (1195713 to LNA); and by the Malcolm M. Marquis, MD, Endowed Fund for Innovation (MEP).

Disclosure: **M.G. Maguire**, Annexon (C), Janssen Research & Development (C), Regenxbio (C); **D.G. Birch**, 4D Molecular Therapeutics (C, F), Applied Genetic Technologies Corp./Beacon (F), Aldyra (C), BlueRock Therapeutics (C), Editas Medicine (C), Nacuity Pharma (S), Ocugen (F), ONL Therapeutics (C), PYC Therapeutics (C, F), SepulBio (C); **J.L. Duncan**, Acucela (F), AGTC X-Linked Retinoschisis Trial (S), Applied Genetic Technologies Corp. (F), Allergan/Abbvie (F), Biogen/Nightstar Therapeutics (F), NEI STEM Trial (S), ProQR (F), PYC Therapeutics (F), Spark Therapeutics Choroideremia Trial (S); **A.R. Ayala**, None; **L.N. Ayton**, Apellis (C), Kiora Pharmaceuticals (C), Novartis (F); **J.K. Cheetham**, Allergan/Abbvie (I); **P. Cheng**, None; **T.A. Durham**, None; **F.L. Ferris**, 4D Molecular Therapeutics (C), Adverum (C), Annexon (C), Apellis (C), Bausch & Lomb (P), Genentech (C), Janssen Research & Development (C), Kodiak (C), Novo Nordisk (C), Roche (C), Regenxbio (C); **C.B. Hoyng**, None; **R.M. Huckfeldt**, Beacon Therapeutics (F), Biogen (F), BlueRock Therapeutics (C), Janssen (C, F), MeiraGTx (F), ProQR (F), Sanofi (C), Spark Therapeutics (F), Sunovion (C); **G.J. Jaffe**, 4D Molecular Therapeutics (C), Adverum (C), Annexon (C), Boehringer Ingelheim (C), Eyepoint

(C), Novartis (C), Regeneron (C), Roche/Genentech (C); **C. Kay**, None; **E.M. Lad**, 4D Molecular Therapeutics (C), Alexion (C, F), Allegro (C), Annexon (C), Apellis (C, F), Aspen Neuroscience (C), Belite Bio (F), BlueRock Therapeutics (C), Boehringer Ingelheim (C, F), Broadwing Bio (C), Complement Therapeutics (C), Galimedix (C), Gemini Therapeutics (F), IVERIC Bio/Astellas (C, F), Janssen (C, F), Kriya Therapeutics (C), LumiThera (C, F), Lutronic Vision (C), Nanoscope Therapeutics (C), Neurotech (F), NGM Biopharmaceuticals (C), Novartis (C), Ocular Therapeutics (C), Osanni Bio (C, I), Perceive Bio (C), Regeneron (C), Retrotope (C), Roche (C, F), Sanofi (C), Thea Laboratoires (C); **B.P. Leroy**, 4D Molecular Therapeutics (C), AAVantgardeBio (C), Akouos (C), Alia Therapeutics (C), Atsena Therapeutics (C), Belite Bio (C), Biogen (C), Coave Therapeutics (C), GenSight Biologics (C), Gyroscope (C), IVERIC Bio (C), Jansen Pharmaceuticals–Johnson & Johnson (C), MeiraGTx (C), Novartis (C), Opus Genetics (C), ProQR Therapeutics (C), Ray Therapeutics (C), Regenxbio (C), Santen (C), Sepul Bio (C), SparingVision (C), Spark Therapeutics (C), SpliceBio (C), Stoke Therapeutics (C), Transine Therapeutics (C), Vedere Bio I & II (C), ViGeneron (C); **W. Liang**, None; **L.S. McDaniel**, None; **M. Melia**, None; **M. Michaelides**, Belite Bio (C), MeiraGTx (C, I, O), Mogrify (C), Saliogen (C); **M.E. Pennesi**, 4D Molecular Therapeutics (C), Adverum (C), Arrowhead Pharmaceuticals (C), Akous (S), Aldebaran (C, I), Applied Genetic Technologies Corp. (C, F), Ascidian (C), Atsena (C, I), Astellas (C), Biogen (I), BlueRock Therapeutics/Opsis (C), Coave (C), ClarisBio (C), Dompe (C), Editas Medicine (C, F), Edigene (C), Endogena (C, I), EnterX (I), Gensight (S), Ingel Therapeutics (C, I), J-Cyte (C), Janssen (C), Kala Therapeutics (C), Kiora (C, I), Nacuity Pharmaceuticals (C, I), Ocugen (C, I), Ora (C), ProQR (C, F), Prime Editing (C), PTC Therapeutics (C), PYC Therapeutics (C), Ray Therapeutics (C), RegenxBio (C), Rejuvitas (C), Reneuron (F), RestoreVision (C), Sanofi (F), Sparing Vision (C), SpliceBio (C), Spotlight Therapeutics (C), Thea (C), Theranexus (C), ZipBio (I); **J.-A. Sahel**, Avista (C, I), Chronolife (I), Cilense (I), Gensight (I), Prophesee (I), Sharpeye (I), Sparing Vision (I), Tenpoint (C, I), Tilak (I), UPMC Enterprises (C), Vegavect (I); **L. Samarakoon**, None

References

1. Ayala A, Cheetham J, Durham T, Maguire M. The importance of natural history studies in inherited retinal diseases. *Cold Spring Harb Perspect*

- Med.* 2023;13(3):a041297, <https://doi.org/10.1101/cshperspect.a041297>.
- Durham TA, Duncan JL, Ayala AR, et al. Tackling the challenges of product development through a collaborative rare disease network: the Foundation Fighting Blindness Consortium. *Transl Vis Sci Technol.* 2021;10(4):23, <https://doi.org/10.1167/tvst.10.4.23>.
 - Kaiserman N, Obolensky A, Banin E, Sharon D. Novel *USH2A* mutations in Israeli patients with retinitis pigmentosa and Usher syndrome type 2. *Arch Ophthalmol.* 2007;125(2):219–224, <https://doi.org/10.1001/archophth.125.2.219>.
 - McGee TL, Seyedahmadi BJ, Sweeney MO, Dryja TP, Berson EL. Novel mutations in the long isoform of the *USH2A* gene in patients with Usher syndrome type II or non-syndromic retinitis pigmentosa. *J Med Genet.* 2010;47(7):499–506, <https://doi.org/10.1136/jmg.2009.075143>.
 - Le Quesne Stabej P, Saihan Z, Rangesh N, et al. Comprehensive sequence analysis of nine Usher syndrome genes in the UK National Collaborative Usher Study. *J Med Genet.* 2012;49(1):27–36, <https://doi.org/10.1136/jmedgenet-2011-100468>.
 - Iannaccone A, Kritchevsky SB, Ciccarelli ML, et al. Kinetics of visual field loss in Usher syndrome type II. *Invest Ophthalmol Vis Sci.* 2004;45(3):784–792, <https://doi.org/10.1167/iovs.03-0906>.
 - Fishman GA, Bozbeyoglu S, Massof RW, Kimberling W. Natural course of visual field loss in patients with Type 2 Usher syndrome. *Retina.* 2007;27(5):601–608, <https://doi.org/10.1097/OI.ae.0000246675.88911.2c>.
 - Sandberg MA, Rosner B, Weigel-DiFranco C, McGee TL, Dryja TP, Berson EL. Disease course in patients with autosomal recessive retinitis pigmentosa due to the *USH2A* gene. *Invest Ophthalmol Vis Sci.* 2008;49(12):5532–5539, <https://doi.org/10.1167/iovs.08-2009>.
 - Jacobson SG, Roman AJ, Aleman TS, et al. Normal central retinal function and structure preserved in retinitis pigmentosa. *Invest Ophthalmol Vis Sci.* 2010;51(2):1079–1085, <https://doi.org/10.1167/iovs.09-4372>.
 - Pierrache LH, Hartel BP, van Wijk E, et al. Visual prognosis in *USH2A*-associated retinitis pigmentosa is worse for patients with Usher syndrome Type IIa than for those with non-syndromic retinitis pigmentosa. *Ophthalmology.* 2016;123(5):1151–1160, <https://doi.org/10.1016/j.opht.2016.01.021>.
 - Duncan JL, Liang W, Maguire MG, et al. Baseline visual field findings in the RUSH2A study: associated factors and correlation with other measures of disease severity. *Am J Ophthalmol.* 2020;219:87–100.
 - Hufnagel RB, Liang W, Duncan JL, et al. Tissue-specific genotype-phenotype correlations among *USH2A*-related disorders in the RUSH2A study. *Hum Mutat.* 2022;43(5):613–624, <https://doi.org/10.1002/humu.24365>.
 - Lad EM, Duncan JL, Liang W, et al. Baseline microperimetry and OCT in the RUSH2A study: structure–function association and correlation with disease severity. *Am J Ophthalmol.* 2022;244:98–116, <https://doi.org/10.1016/j.ajo.2022.08.013>.
 - 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(11):9, <https://doi.org/10.1167/tvst.9.11.9>.
 - Iannaccone A, Brewer CC, Cheng P, et al. Auditory and olfactory findings in patients with *USH2A*-related retinal degeneration—findings at baseline from the rate of progression in *USH2A*-related retinal degeneration natural history study (RUSH2A). *Am J Med Genet A.* 2021;185(12):3717–3727, <https://doi.org/10.1002/ajmg.a.62437>.
 - Birch DG, Samarakoon L, Melia M, et al. The RUSH2A study: dark-adapted visual fields in patients with retinal degeneration associated with biallelic variants in the *USH2A* gene. *Invest Ophthalmol Vis Sci.* 2022;63(3):17, <https://doi.org/10.1167/iovs.63.3.17>.
 - Duncan JL, Cheng P, Maguire MG, et al. Static perimetry in the rate of progression in *USH2A*-related retinal degeneration (RUSH2A) study: assessment through 2 years. *Am J Ophthalmol.* 2023;250:103–110, <https://doi.org/10.1016/j.ajo.2023.02.002>.
 - Duncan JL, Liang W, Maguire MG, et al. Change in cone structure over 24 months in *USH2A*-related retinal degeneration. *Am J Ophthalmol.* 2023;252:77–93, <https://doi.org/10.1016/j.ajo.2023.03.006>.
 - Beck RW, Moke PS, Turpin AH, et al. A computerized method of visual acuity testing: adaptation of the early treatment of diabetic retinopathy study testing protocol. *Am J Ophthalmol.* 2003;135(2):194–205, [https://doi.org/10.1016/s0002-9394\(02\)01825-1](https://doi.org/10.1016/s0002-9394(02)01825-1).
 - Evans J. *Straightforward Statistics for the Behavioral Sciences.* Pacific Grove, CA: Brooks/Cole Publishing; 1996:146.
 - Huber PJ. Robust regression: asymptotics, conjectures and Monte Carlo. *Ann Stat.* 1973;1:799–821.

22. Chen C. Robust regression and outlier detection with the ROBUSTREG procedure, Paper 265026. Cary, NC: SAS Institute; 2002.
23. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*. 1986;1(8476):307–310.
24. Blackhurst DW, Maguire MG. Reproducibility of refraction and visual acuity measurement under a standard protocol. The Macular Photocoagulation Study Group. *Retina*. 1989;9(3):163–169.
25. Csaky K, Ferris F, 3rd, Chew EY, Nair P, Cheetham JK, Duncan JL. Report from the NEI/FDA endpoints workshop on age-related macular degeneration and inherited retinal diseases. *Invest Ophthalmol Vis Sci*. 2017;58(9):3456–3463, <https://doi.org/10.1167/iovs.17-22339>.
26. Csaky KG, Richman EA, Ferris FL. Report from the NEI/FDA Ophthalmic Clinical Trial Design and Endpoints Symposium. *Invest Ophthalmol Vis Sci*. 2008;49(2):479–489, <https://doi.org/10.1167/iovs.07-1132>.
27. Weinreb RN, Kaufman PL. The glaucoma research community and FDA look to the future: a report from the NEI/FDA CDER Glaucoma Clinical Trial Design and Endpoints Symposium. *Invest Ophthalmol Vis Sci*. 2009;50(4):1497–1505, <https://doi.org/10.1167/iovs.08-2843>.
28. Levin LA, Sengupta M, Balcer LJ, Kupersmith MJ, Miller NR. Report from the National Eye Institute Workshop on Neuro-Ophthalmic Disease Clinical Trial Endpoints: Optic Neuropathies. *Invest Ophthalmol Vis Sci*. 2021;62(14):30, <https://doi.org/10.1167/iovs.62.14.30>.
29. Hood DC, Ramachandran R, Holopigian K, Lazow M, Birch DG, Greenstein VC. Method for deriving visual field boundaries from OCT scans of patients with retinitis pigmentosa. *Biomed Opt Express*. 2011;2(5):1106–1114.
30. 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(8):N51–N56, <https://doi.org/10.1088/0967-3334/28/8/n02>.
31. Klein M, Birch DG. Psychophysical assessment of low visual function in patients with retinal degenerative diseases (RDDs) with the Diagnosys full-field stimulus threshold (D-FST). *Doc Ophthalmol*. 2009;119(3):217–224, <https://doi.org/10.1007/s10633-009-9204-7>.
32. Aleman TS, Uyhazi KE, Serrano LW, et al. *RDH12* mutations cause a severe retinal degeneration with relatively spared rod function. *Invest Ophthalmol Vis Sci*. 2018;59(12):5225–5236, <https://doi.org/10.1167/iovs.18-24708>.
33. 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(10097):849–860, [https://doi.org/10.1016/S0140-6736\(17\)31868-8](https://doi.org/10.1016/S0140-6736(17)31868-8).
34. European Medicines Agency. Luxturna. Available at: <https://www.ema.europa.eu/en/medicines/human/EPAR/luxturna>. Accessed September 18, 2024.
35. Wu Z, Crabb DP, Chauhan BC, Crowston JG, Medeiros FA. Improving the feasibility of glaucoma clinical trials using trend-based visual field progression endpoints. *Ophthalmol Glaucoma*. 2019;2(2):72–77, <https://doi.org/10.1016/j.ogla.2019.01.004>.
36. Montesano G, Garway-Heath DF, Rabiolo A, De Moraes CG, Ometto G, Crabb DP. Validating trend-based end points for neuroprotection trials in glaucoma. *Transl Vis Sci Technol*. 2023;12(10):20, <https://doi.org/10.1167/tvst.12.10.20>.
37. Altman DG, Royston P. The cost of dichotomising continuous variables. *BMJ*. 2006;332(7549):1080, <https://doi.org/10.1136/bmj.332.7549.1080>.
38. Massof RW, Stelmack JA. Interpretation of low-vision rehabilitation outcome measures. *Optom Vis Sci*. 2013;90(8):788–798, <https://doi.org/10.1097/oxp.000000000000011>.