IRD CLINICAL TRIAL ENDPOINTS: Lessons Learned

Choosing the right outcome measures for inherited retinal disease trials requires matching endpoints to disease biology, therapeutic mechanism, and patient experience.

By Melissa Yuan, MD, and Ahmad Al Moujahed, MD, PhD

Conventional ophthalmic endpoints using BCVA were developed for conditions such as cataract surgery, AMD, and diabetic retinopathy, for which large sample sizes and central BCVA changes are readily captured. Inherited retinal diseases (IRDs) challenge these conventional endpoints because patients with conditions such as retinitis pigmentosa (RP) may retain a VA of 20/20 for years while losing peripheral and night vision—changes that BCVA cannot detect. At the same time, patients with end-stage disease may already be at the measurement threshold, leaving no room to capture further visual decline.[1,2] Phenotypic heterogeneity, slow progression, and small patient populations compound these difficulties.[3]

The expanding diversity of potential therapeutic strategies means endpoints must be matched to both disease phenotype and mechanism of intervention.[4] With numerous candidates advancing through clinical development, endpoint selection has emerged as a critical determinant of trial success or failure. For example, the phase 3 STAR trial of timrepigene emparvovec (BIIB111, Biogen) for choroideremia investigated a primary endpoint of a gain of 15 ETDRS letters or more, which was borrowed from wet AMD trials where anti-VEGF therapy routinely produces such gains. Although a meaningful proportion of treated patients showed 10-letter gains versus controls, the rigid primary endpoint did not capture this signal, given the disease's slow progression.[5] Borrowing endpoints from other retinal diseases without adapting them to the natural history of the specific IRD risks masking a genuine treatment effect.[6]

ESTABLISHED CLINICAL ENDPOINTS

BCVA, a component of almost all IRD trials, offers regulatory familiarity but alone may be insufficient.[1,4,7] In the

pivotal trial for voretigene neparvovec-rzyl (Luxturna, Spark Therapeutics) for *RPE65*-associated retinal dystrophy, BCVA did not reach statistical significance between arms; efficacy was instead demonstrated through the multi-luminance mobility test (MLMT), which measured navigation ability through a maze and past obstacles in dim light—a clinically meaningful endpoint.

Electronic alternatives such as the Freiburg Visual Acuity Test extend measurement into the hand-motions range, serving as the primary endpoint in the phase 2b RESTORE trial of MCO-010 (Nanoscope) for advanced RP, in which patients' vision was so poor, it was "off chart."[8]

BCVA is most responsive when the disease disrupts central function early (as in achromatopsia or Leber congenital amaurosis [LCA] genotypes, including *GUCY2D*, *CEP290*, and *RPE65*), but it tells little about progression in rod-cone dystrophies where patients may continue to read 20/25 while their peripheral fields collapse.[4]

Contrast sensitivity remains underused in IRD trials. The standard Pelli-Robson chart samples a limited range and bottoms out quickly. Newer automated platforms such as the quick contrast sensitivity function have shown stronger correlation with patient-reported quality of life than BCVA.[9]

The MLMT is the only patient-focused outcome that has been approved as a primary endpoint for an IRD gene therapy trial. Its strength lies in directly measuring the patient's ability to move safely through an environment. However, the MLMT is labor-intensive, requires dedicated physical space and personnel, and is rarely used outside trial settings, complicating long-term monitoring.[1] Newer alternatives such as virtual reality mobility assessments may address these limitations.[10] Researchers have also expanded on standardized performance-based tests, including physical and virtual reality-based paradigms developed at the StreetLab (now at UPMC Vision Institute) that track gaze, gait, and head motion during navigation tasks across multiple lighting levels; these tests have been adopted for the RUSH2A follow-up studies (NCT03146078).[11]

Visual field testing is conceptually appealing for rod-cone dystrophies in which peripheral vision is lost, but standard automated perimetry has high test-retest variability in scotomatous regions.[12] Microperimetry, with real-time fundus tracking, partly addresses this and is designated as an efficacy endpoint in the phase 2/3 VISTA trial of laru-zova (AGTC-501, Beacon Therapeutics) for *RPGR*-related X-linked RP (XLRP).[3] A key statistical challenge is multiplicity: When sensitivity is tested at dozens of loci, the probability of false-positive improvements increases sharply.[13] The FDA has required a mean improvement of at least 7 dB in at least five prespecified loci, but this conservative approach contributed to the XIRIUS phase 2/3 study of cotoretigene toliparvovec (BIIB112, Biogen) for XLRP failing its primary endpoint despite apparent treatment signals.[14] An alternative binomial distribution approach requiring at least seven unspecified loci with at least 7 dB improvement across a 68-locus grid, as applied to the phase 2 SKYLINE trial data of laru-zova in XLRP, may offer a more balanced approach while maintaining the false-positive rate near 5% without the need for prespecification.[13]

For trials targeting rod-specific rescue, scotopic microperimetry and two-color dark-adapted perimetry can isolate rod-mediated versus cone contributions, although both add time and complexity.[4] Two-year data from the RUSH2A study demonstrated that quantitative static perimetry measures declined significantly in eyes with *USH2A*-related retinal degeneration while central measures declined more slowly.[15] Four-year results found that testing at functional transition points, which are loci at the boundary of the ellipsoid zone (EZ), was particularly sensitive and that rates of change were substantially more efficient than threshold-based proportions, often requiring half the sample size for equivalent statistical power.[16]

KEY TAKEAWAYS

- BCVA as a trial endpoint is most responsive when the disease disrupts central function early but tells little about progression in rod-cone dystrophies.
- The multi-luminance mobility test, Freiburg Visual Acuity Test, low-luminance visual acuity, full-field stimulus threshold testing, fundus autofluorescence, OCT-derived metrics, composite endpoints, and patient-reported outcome measures may all serve as more accurate endpoints in specific inherited retinal disease trials.
- With more than two dozen gene therapy trials in phase 2 or 3, endpoint strategy is integral to whether an effective therapy succeeds or fails on paper.

EMERGING ENDPOINTS

Low-luminance visual acuity (LLVA) adds a neutral density filter to simulate dim conditions, revealing functional deficits that are invisible under standard lighting. Many patients with choroideremia or *RPGR*-related RP report difficulty in dim environments before their standard acuity begins to decline.[17] LLVA has been approved by the FDA as an efficacy endpoint in the VISTA trial for *RPGR*-related XLRP.[18]

Full-field stimulus threshold testing (FST) is feasible in patients with nystagmus or eccentric fixation who cannot perform standard perimetry. Using chromatic stimuli can isolate rod versus cone contributions.[19] Japanese regulators accepted FST as the primary endpoint for voretigene

TABLE. OVERVIEW OF SELECT CLINICAL TRIAL ENDPOINTS FOR INHERITED RETINAL DISEASE

Endpoint	Principal Strengths	Principal Limitations	Regulatory Precedent
CENTRAL VISUAL FUNCTION			
BCVA	• Widely validated • Standardized charts	• Insensitive to peripheral/low-light deficits • Ceiling/floor effects	Accepted primary endpoint across multiple inherited retinal disease programs
Contrast sensitivity	• Detects contrast-based functional losses that BCVA misses (eg, difficulty in dim light or low-contrast environments) • Covers range of spatial frequencies	• Moderate inter-rater variability • Not standardized across devices	Secondary measure in select trials
LIGHT SENSITIVITY AND VISUAL FIELDS			
Static and kinetic perimetry	• Semi-automated • Standard statistical analysis • Established in glaucoma	• Insensitive at severely reduced baseline • Limited ability to detect small changes	Secondary endpoint in *RPGR* and *RPE65* trials
Microperimetry (mesopic and scotopic)	• Dark-adapted protocols assess rod function • Targeted stimulus on retinal loci	• Substantial learning curve • High test-retest and inter-rater variability	Secondary/exploratory in select gene therapy programs
Full-field stimulus threshold	• Practical post-treatment measure • Adaptable across disease stages	• Needs further validation as standalone primary • Overall sensitivity only	Primary in Japan (*RPE65*); secondary elsewhere
REAL-WORLD FUNCTIONAL PERFORMANCE			
Multi-luminance mobility test	• Reflects real-world visual performance holistically • Accepted as pivotal primary endpoint	• Labor-intensive • Observer-dependent • Not suited for routine clinical use	Primary endpoint in voretigene neparvovec-rzyl (Luxturna, Spark) pivotal trial
STRUCTURAL/ANATOMICAL IMAGING			
Ellipsoid zone and fundus autofluorescence imaging	• Objective • Correlates with phenotypic progression • High measurement precision	• Indirect link to visual function • Specialized equipment and expertise needed	Primary anatomical outcome in geographic atrophy and Stargardt trials

neparvovec-rzyl approval, supported by data showing concordance between FST thresholds and mobility performance.[19] In the RUSH2A study, FST showed worsening throughout the wide range of baseline values, with 47% of eyes exceeding the coefficient of repeatability at 4 years, making FST one of the most sensitive tests for detecting change.[16] However, the clinical effect of FST changes in patients with good central vision remains to be established.[16] Furthermore, a multi-stakeholder consensus concluded that FST is not yet ready as a primary endpoint, requiring further standardization and stronger linkage to patient-centered measures.[11]

Structural imaging has long played a supporting role in IRD trials, but regulatory acceptance of fundus autofluorescence (FAF)-measured atrophy as a primary endpoint in geographic atrophy trials has opened the door for imaging-based outcomes.[20] For IRDs, the most promising OCT-derived metrics are the EZ width and area, serving as surrogates for surviving photoreceptor territory. EZ measurements track disease progression with reasonable sensitivity.[21] However, the RUSH2A study found that only 33% of eyes had sufficient baseline EZ area (≥ 3 mm^2) to detect measurable change, and among those eyes, the standardized rate of change was lower than for most functional measures. These findings led to the recommendation that clinical trials for *USH2A* should require a minimum EZ area of at least 3 mm^2 at enrollment to ensure the endpoint can capture progression.[16] FAF provides complementary data, with loss of signal indicating retinal pigment epithelium loss and preserved autofluorescence area tracking progression in choroideremia and RP.[22]

Composite endpoints combining structural and functional measures may increase statistical sensitivity. In *ABCA4*-associated Stargardt disease, the most sensitive composite performed best without BCVA, which contributed more noise than signal.[23] Machine learning tools are also being applied to combine imaging and functional datasets, with early work in LCA suggesting that algorithmic

approaches can identify localized treatment potential that no single clinical test would.[24]

Patient-reported outcome measures have been limited to secondary endpoints in IRD trials due to concerns about subjectivity and variability. Newer IRD-specific instruments such as the Michigan Retinal Degeneration Questionnaire and the ViSIO-PRO tools may provide validated, disease-specific assessments.[25,26] While objective assessments detect biological treatment effects, changes in these measures may not necessarily translate into improvements patients perceive as meaningful, underscoring the importance of incorporating patient perspectives.

PRACTICAL CONSIDERATIONS

For investigators designing IRD trials, several principles emerge. Endpoint selection should be driven by the pathophysiology of the genetic condition: A rod-cone dystrophy preserving central vision requires different outcome measures than a macular dystrophy. The therapeutic mechanism should also play a role. Gene augmentation aiming to halt photoreceptor loss may be best served by a structural endpoint or functional mobility assessments for rod-mediated function in rod-cone dystrophies. An optogenetic therapy restoring light sensitivity to surviving inner retinal neurons may require a functional mobility assessment to capture the qualitative shift in visual capability. Small patient populations may necessitate surrogate endpoints offering greater sensitivity.

Across the trial landscape, endpoint choices have been tailored to individual genotypes (Table). In *RPE65*-associated LCA, the approved therapy validated the complementary use of both mobility and psychophysical endpoints. For choroideremia, multiple completed phase 1/2 trials relying on microperimetry, FAF, and BCVA have not demonstrated statistically significant treatment effects, fueling debate about whether these endpoints are sufficiently sensitive for slowly progressive diseases.[27,28] For achromatopsia, the stationary nature and the primacy of photophobia have led investigators to incorporate light sensitivity measures as functional outcomes, including palpebral aperture quantification and visual photosensitivity thresholds.[4]

ALIGNING ENDPOINTS WITH LIVED EXPERIENCE

With more than two dozen gene therapy trials in phase 2 or 3, endpoint strategy is integral to whether an effective therapy succeeds or fails on paper. The retina specialist's role in this process is becoming more important, as familiarity with the clinical nuances, the practical realities of administering and interpreting these tests, and the lived experience of patients all inform which measurements will matter. As the regulatory landscape evolves, the goal is not simply more endpoints but better alignment between what we measure and what patients gain from treatment. ■

1. Thirunavukarasu AJ, Raji S, Cehajic Kapetanovic J. Visualising treatment effects in low-vision settings: proven and potential endpoints for clinical trials of inherited retinal disease therapies [published online ahead of print August 7, 2025]. *Gene Ther*.
2. Simunovic MP, Grigg JR, Mahroo OA. Vision at the limits: Absolute threshold, visual function, and outcomes in clinical trials. *Surv Ophthalmol*. 2022;67(4):1270-1286.
3. Igoe JM, Lam BL, Gregori NZ. Update on clinical trial endpoints in gene therapy trials for inherited retinal diseases. *J Clin Med*. 2024;13(18):5512.
4. Georgiou M, Fujinami K, Michaelides M. Inherited retinal diseases: Therapeutics, clinical trials and end points–A review. *Clin Exp Ophthalmol*. 2021;49(3):270-288.
5. MacLaren RE, Fischer MD, Gow JA, et al. Subretinal timrepigene emparvovec in adult men with choroideremia: a randomized phase 3 trial. *Nat Med*. 2023;29(10):2464-2472.
6. Abdalla Elsayed MEA, Cehajic-Kepetanovic J, MacLaren RE. Gene therapy for choroideremia: progress, potential and pitfalls. *Expert Opin Biol Ther*. 2025;25(3):257-263.
7. Schmetterer L, Scholl H, Garhöfer G, et al. Endpoints for clinical trials in ophthalmology. *Prog Retin Eye Res*. 2023;97:101160.
8. Boyer DS, Bergstrom L, Emanuelli A, et al. Efficacy and safety of MCO-010 optogenetic therapy for vision restoration in patients with severe vision loss due to retinitis pigmentosa: A phase 2b randomized, sham-controlled, multi-center, multi-dose, double-masked clinical trial (RESTORE). *Invest Ophthalmol Vis Sci*. 2023;64(8):5443-5443.
9. Vingopoulos F, Bannerman A, Zhou P, et al. Towards the validation of quantitative contrast sensitivity as a clinical endpoint: correlations with vision-related quality of life in bilateral AMD. *Br J Ophthalmol*. 2024;108(6):846-851.
10. Aleman TS, Miller AJ, Maguire KH, et al. A virtual reality orientation and mobility test for inherited retinal degenerations: testing a proof-of-concept after gene therapy. *Clin Ophthalmol*. 2021;15:939-952.
11. Ocular Diseases Forum 3 - Forum for Collaborative Research. Accessed March 24, 2026. www.forumresearch.org/ocular/meetingsocculardisease/1866-ocular-diseases-forum-3
12. Gardiner SK, Swanson WH, Goren D, Mansberger SL, Demirel S. Assessment of the reliability of standard automated perimetry in regions of glaucomatous damage. *Ophthalmology*. 2014;121(7):1359-1369.
13. Yaghy A, Birch DG, Hwang Y, et al. Addressing multiplicity in retinal sensitivity analysis: an alternative approach to assessing gene therapy efficacy in inherited retinal diseases. *Trans Vis Sci Tech*. 2025;14(3):25.
14. Lam BL, Pennesi ME, Kay CN, et al. Assessment of visual function with cotoretigene toliparvovec in x-linked retinitis pigmentosa in the randomized XIRIUS phase 2/3 study. *Ophthalmology*. 2024;131(9):1083-1093.
15. 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.
16. Maguire MG, Birch DG, Duncan JL, et al. Endpoints and design for clinical trials in *USH2A*-related retinal degeneration: results and recommendations from the RUSH2A natural history study. *Trans Vis Sci Tech*. 2024;13(10):15.
17. Wood LJ, Jolly JK, Buckley TM, Josan AS, MacLaren RE. Low luminance visual acuity as a clinical measure and clinical trial outcome measure: a scoping review. *Ophthalmic Physiol Opt*. 2021;41(2):213-223.
18. A study of laru-zova in participants with X-linked retinitis pigmentosa (VISTA). ClinicalTrials.gov. Accessed May 21, 2026. clinicaltrials.gov/study/NCT04850118
19. Shi LF, Hall AJ, Thompson DA. Full-field stimulus threshold testing: a scoping review of current practice. *Eye (Lond)*. 2024;38(1):33-53.
20. Mai J, Reiter GS, Riedl S, et al. Quantitative comparison of automated OCT and conventional FAF-based geographic atrophy measurements in the phase 3 OAKS/DERBY trials. *Sci Rep*. 2024;14(1):20531.
21. Zada M, Cornish EE, Fraser CL, Jamieson RV, Grigg JR. Natural history and clinical biomarkers of progression in X-linked retinitis pigmentosa: a systematic review. *Acta Ophthalmologica*. 2021;99(5):499-510.
22. Pichi F, Abboud EB, Ghazi NG, Khan AO. Fundus autofluorescence imaging in hereditary retinal diseases. *Acta Ophthalmologica*. 2018;96(5):e549-e561.
23. Lambertus S, Bax NM, Fakin A, et al. Highly sensitive measurements of disease progression in rare disorders: Developing and validating a multimodal model of retinal degeneration in Stargardt disease. *Plos One*. 2017;12(3):e0174020.
24. Sumaroka A, Garafalo AV, Semenov EP, et al. Treatment potential for macular cone vision in Leber congenital amaurosis due to *CEP290* or *NPHP5* mutations: predictions from artificial intelligence. *Invest Ophthalmol Vis Sci*. 2019;60(7):2551-2562.
25. Fischer MD, Patalano F, Naujoks C, et al. Psychometric validation of the ViSIO-PRO and ViSIO-ObsRO in retinitis pigmentosa and Leber congenital amaurosis. *Ophthalmol Ther*. 2023;12(2):1359-1386.
26. Lacy GD, Abalem MF, Andrews CA, et al. The Michigan retinal degeneration questionnaire: a patient-reported outcome instrument for inherited retinal degenerations. *Am J Ophthalmol*. 2021;222:60-68.
27. Lam BL, Davis JL, Gregori NZ, et al. Choroideremia gene therapy phase 2 clinical trial: 24-month results. *Am J Ophthalmol*. 2019;197:65-73.
28. Shen LL, Ahluwalia A, Sun M, Young BK, Nardini HKG, Priore LVD. Long-term natural history of visual acuity in eyes with choroideremia: a systematic review and meta-analysis of data from 1004 individual eyes. *British J Ophthalmol*. 2021;105(2):271-278.

MELISSA YUAN, MD

- Vitreoretinal Surgery Fellow, Retina Service, Byers Eye Institute, Stanford University, Palo Alto, California
- melissayyuan@gmail.com
- Financial disclosure: None

AHMAD AL MOUJAHED, MD, PHD

- Assistant Professor of Ophthalmology, Vitreoretinal Surgeon, Inherited Retinal Degenerations Specialist, Byers Eye Institute, Stanford University, Palo Alto, California
- ahmadm@stanford.edu
- Financial disclosure: None