Published in final edited form as:

Doc Ophthalmol. 2009 December ; 119(3): 217–224. doi:10.1007/s10633-009-9204-7.

Psychophysical assessment of low visual function in patients with retinal degenerative diseases (RDDs) with the Diagnosys full-field stimulus threshold (D-FST)

M. Klein and

Retina Foundation of the Southwest, 9900 N. Central Expressway, Suite 400, Dallas, TX 75231, USA

D. G. Birch

Retina Foundation of the Southwest, 9900 N. Central Expressway, Suite 400, Dallas, TX 75231, USA. Department of Ophthalmology, UT Southwestern Medical Center, Dallas, TX, USA

Abstract

To determine whether the Diagnosys full-field stimulus threshold (D-FST) is a valid, sensitive and repeatable psychophysical method of measuring and following visual function in low-vision subjects. Fifty-three affected eyes of 42 subjects with severe retinal degenerative diseases (RDDs) were tested with achromatic stimuli on the D-FST. Included were subjects who were either unable to perform a static perimetric field or had non-detectable or sub-microvolt electroretinograms (ERGs). A subset of 21 eyes of 17 subjects was tested on both the D-FST and the FST2, a previous established full-field threshold test. Seven eyes of 7 normal control subjects were tested on both the D-FST and the FST2. Results for the two methods were compared with the Bland–Altman test. On the D-FST, a threshold could successfully be determined for 13 of 14 eyes with light perception (LP) only (median $0.9 \pm 1.4 \log \text{cd/m}^2$), and all eyes determined to be counting fingers (CF; median $0.3 \pm 1.8 \log \text{cd/m}^2$). The median full-field threshold for the normal controls was $-4.3 \pm 0.6 \log \text{cd/m}^2$ on the D-FST and $-4.8 \pm 0.9 \log \text{cd/m}^2$ on the FST2. The D-FST offers a commercially available method with a robust psychophysical algorithm and is a useful tool for following visual function in low vision subjects.

Keywords

Psychophysics; Treatment trials; Retina; LCA; RP; CRD

Introduction

Advances in subretinal cell therapy [1], subretinal delivery of corrective genes [2–5] and visual prostheses [6, 7] require the participation of patients with minimal visual function in clinical trials. However, once vision drops below a level that can be measured with common visual function tests like visual field perimetry or electroretinography (ERG) and visual acuity falls to the level of light perception (LP), there is no commonly accepted method

available to follow levels of progression or improvement. The development of the full-field stimulus test (FST) [8] created a psychophysical method to measure the illuminances necessary to stimulate the most sensitive parts of the retina and thus deliver a quantifiable threshold that could be assessed before and after intervention even in subjects with extremely low vision. The most sensitive area can be measured without knowing its exact location, which is especially useful since areas of improvement in trials utilizing subretinal injections often are very localized and patchy, making it difficult to pick up in functional measurements across large areas of retina (e.g. full-field electroretinography). As an additional benefit, this method does not require fixation, making it useful even for subjects with better vision who are unable to perform static visual field perimetry due to conditions like nystagmus. The further development of the method created the FST2 [9], moving the test from a customized hardware base [8] to a software level on the Espion ColorDome™ LED full-field stimulator (Diagnosys LLC, Lowell, MA) [9]. This software, however, is not commercially available. With the inclusion of the newly developed Diagnosys full-field stimulus test (D-FST) in the software library of the Espion system, a technique for obtaining full-field thresholds has become commercially available and could potentially be useful in both scientific and clinical settings. Since there are, however, differences in stimulus parameters and the psychophysical protocol between the D-FST and the FST2, it is necessary to evaluate the D-FST in comparison with the established FST2.

In this study, we evaluate patients with minimal visual function on the D-FST to determine validity, sensitivity and repeatability and, in a subset of patients, compared the D-FST to the FST2.

Methods

Subjects

We recruited 42 patients between 5 and 84 years of age who were severely affected by retinal degenerative diseases (RDDs; see Table 1 for breakdown of demographics by condition) and 7 normal subjects between the age of 23 and 52 years of age. Thirty-eight patients were recruited because they had non-detectable rod ERG amplitudes ($<2.0 \mu\text{V}$) and non-detectable ($<0.1 \mu\text{V}$) or sub-microvolt ($<1.0 \mu\text{V}$) cone ERG amplitudes. These patients also could not be tested with static perimetry. We recruited 4 additional patients retaining measurable rod and/or cone ERGs because they were unable to reliably read the eye chart or perform Humphrey static perimetry due to lack of fixation. Normal subjects had visual acuity of $\leq 0.1 \log\text{MAR}$, normal eye exams and normal full-field ERGs. Monocular testing was conducted to evaluate the remaining visual function in each eye separately, as is often the case in clinical trials where one eye serves as control.

All subjects gave informed consent before testing and were treated in adherence to the tenets of the Declaration of Helsinki.

D-FST

Fifty-three affected eyes of the 42 subjects with RDDs and 7 eyes of 7 normal subjects were tested. Subjects were dilated by topical application of tropicamide 1% and phenylephrine

hydrochloride 2.5% and dark-adapted for 30 min before testing. Subjects were tested monocularly and had the fellow eye patched before testing began. Twenty-four subjects (17 patients, 7 normal subjects) returned for testing on at least three occasions.

The test utilized the Diagnosys Espion system with the ColorDome™ LED full-field stimulator (Diagnosys LLC, Lowell, MA) to create achromatic (65,000 K, a preprogrammed setting in the Espion system) full-field stimuli. In the test setup, a 0-decibel (dB) reference point is chosen, which for our study was 0.1 cd.s/m² (25 cd/m² presented for 4 ms). The range of decibels available for the test varied from -75 dB (dim) to +15 dB (bright), with 10 dB equaling 1 log unit. The tester then defined a starting value to initiate testing. The proprietary program utilized a forced-choice strategy to test within a range of 10 dB around the start value. If the threshold was not found within that 10 dB range, the program shifted that 10 dB range up or down until a threshold was reached. A two parameter Weibull function [10] was used to calculate the actual threshold while taking into account false positives and false negatives.

A two-button box with one button covered with felt (positive answer) and a second, smooth button (negative answer) served as patient interface. Subjects were presented with an auditory cue followed by a stimulus and had to indicate by button-push if they perceived the presented stimulus or not. Subjects had 5 s to make a decision before the program timed-out and presented the next auditory cue. Each run required ~40 trials before a threshold was reached. After a short training run to introduce the subject to the test, three runs were recorded on each subject. Typically, a test session on one eye took ~8 min to complete.

FST2

A subset of 21 eyes of 17 subjects with severe RDDs and 7 eyes of 7 normal subjects were also tested on the same day on the FST2 with a 5 min break between sessions. Subjects were dilated by topical application of tropicamide 1% and phenylephrine hydrochloride 2.5% and dark-adapted for 30 min before testing. Subjects were tested monocularly and had the fellow eye patched before testing began.

The test utilized the Diagnosys Espion system with the ColorDome™ LED full-field stimulator (Diagnosys LLC, Lowell, MA) to create achromatic (65,000 K) full-field stimuli. The 0-dB illuminance is fixed to 0.74 cd.s/m² (3.7 cd/m² presented for 200 ms) and cannot be changed by the user. The range of decibels available for the test ranged from 65 (dim) to -25 dB (bright), with 10 dB equaling 1 log unit.

The FST2 followed a 4-2 staircase method, similar to the full threshold strategy of the Humphrey perimeter [9]. A start value for the staircase was defined by the tester. Subjects were told when the test started and then had to indicate by pushing a button whenever they perceived a stimulus. Each run required ~14 trials before a threshold was reached. After a short training run to introduce the subject to the test, three runs were recorded on each subject. Each eye required ~5 min for testing.

Statistical methods

We compared the tests performed on both the D-FST and the FST2 in the same eye and on the same day in the 21 eye subset of subjects with RDDs and the 7 eyes of normal subjects. To account for flash duration in accordance with Bloch's law (reciprocal relationship between time and intensity) [11], we converted thresholds to luminance (cd/m^2) for better comparison. For the D-FST, the conversion formula was

$$\log L = \log L_{0,\text{D-FST}} + S/10$$

For the FST2, the conversion formula was

$$\log L = \log L_{0,\text{FST2}} - S/10$$

where L is the calculated luminance in cd/m^2 , L_0 is the reference luminance in cd/m^2 and S is the sensitivity (threshold) in dB. A statistical method (Bland–Altman) was used to quantify agreement between the two methods of sensitivity assessment [12]. The difference in sensitivity between the two methods was compared to the mean sensitivity of the two methods.

Test/retest variability of the D-FST

We evaluated the test/retest variability of the D-FST in 27 eyes from 24 subjects. Using a Bland–Altman analysis, we compared the first run to the second run, the first run to the third run and the second run to the third run. The differences between runs were plotted against the mean of the runs, resulting in three comparisons per subject.

Results

Thresholds were successfully obtained with the D-FST for a wide range of ages (5 to 84 years of age). Results from a 5-year-old child with Leber congenital amaurosis (LCA; Fig. 1a) show probability plotted against stimulus intensity. At -32 dB ($-1.8 \log \text{cd/m}^2$), the child failed to detect any flashes. Above -18 dB ($-0.4 \log \text{cd/m}^2$), the patient detected all flashes. Between these two values, the probability of detection rose rapidly. The solid line is a Weibull function fit to the data (solid triangles), with the threshold (midway between floor and ceiling) indicated as a solid rectangle. Other patients failed to respond 100% of the time to intense stimuli (false negatives) or gave occasional spurious responses to dim stimuli (false positives). The Weibull function corrects for these errors by appropriately adjusting the floor, ceiling and threshold (Fig. 1b), with the threshold being equal to the 50% point between the floor and the ceiling of the function.

We were able to obtain reliable thresholds on 51 of the 53 severely affected eyes and 7 of the 7 normal eyes tested with the D-FST. This included 13 of the 14 LP eyes (median $0.9 \pm 1.4 \log \text{cd/m}^2$) and 6 of the 6 counting fingers (CF) eyes (median $0.3 \pm 1.8 \log \text{cd/m}^2$). One eye with no light perception (NLP) and one eye with LP on standard clinical assessment were

unable to perceive the most intense stimulus on the D-FST. All 21 eyes of the subset tested with both methods reached a measurable threshold. Medians, quartiles and full ranges of thresholds for each diagnostic category are shown on both the D-FST and the FST2 in Fig. 2. Median thresholds of eyes of normal subjects tested with both the D-FST and the FST2 were $-4.3 \log \text{ cd/m}^2$ and $-4.8 \log \text{ cd/m}^2$, respectively. Eyes of the RDD subgroup tested on both systems showed an elevation above median normal of 3.8 log units on the D-FST and 3.4 log units on the FST2. Compared to the median normal subject, patients with cone-rod dystrophy (CRD) were elevated the least (0.8 log units on the D-FST and by 1.6 log units on the FST2). The most elevated group was LCA, where the average elevation was 4.6 log units on the D-FST and 4.7 log units on the FST2.

A comparison of the two methods showed good overall correlation ($r = 0.91$; $P < 0.001$; Fig. 3a). A Bland–Altman analysis showed a 0.4 bias against the D-FST, with a 95% confidence range covering an interval between -1.5 and $2.3 \log \text{ cd/m}^2$ (Fig. 3b). Test/retest variability for the D-FST as determined by Bland–Altman test (Fig. 4) was $\pm 0.3 \log \text{ cd/m}^2$.

Multiple examinations over at least 12 months are shown for three patients in Fig. 5. Patient #8093 is a 57-year-old male with retinitis pigmentosa (RP) with bare light perception. He showed a transient increase in threshold at 4 months possibly associated with an episode of viral conjunctivitis. Over the 18 months of follow-up, threshold increased by $0.6 \log \text{ cd/m}^2$, a value well above the repeat variability of the test. Patients #8095 (female, age 41) and #8218 (female, age 67) also had RP with bare light perception. Their thresholds did not change significantly throughout the course of follow-up visits.

Discussion

The results of this study confirm that the D-FST, like the FST2 [9], provides a reliable measurement of retinal sensitivity in severely affected subjects with RDDs. Since patients were not in the end-stage of their diseases, subjects with retinal diseases primarily affecting rods had higher elevations in their threshold than subjects with diseases primarily affecting cones. Thus, subjects with CRD had thresholds close to normal, while subjects with advanced RP were elevated by an average of 4.1 log units. As patients approach end-stage disease, these differences among diagnostic groups are likely to lessen. While subjects with CRD had close to normal thresholds, their difficulties with traditional clinical measurements like static perimetry due to nystagmus make the D-FST a useful measure of their visual function since steady fixation is not important for this method [8]. The brief test duration and simple requirements also facilitate testing young children who may lack the uninterrupted attention span to perform perimetry and the elderly with health problems that prevent them from remaining stationary for prolonged periods of time.

The test/retest variability of $\pm 0.3 \log \text{ cd/m}^2$ is comparable to the $\pm 0.28 \log \text{ cd/m}^2$ value reported for the FST2 [9] and values for more traditional tests of dark-adaption thresholds such as the Goldmann-Weekers or the SST-1 [13]. This reasonable repeat variability makes it feasible to follow changes in retinal sensitivity in longitudinal studies (Fig. 5), which could be useful for judging small improvements in phase 1 treatment trials that involve severely affected patients.

The comparison of the D-FST with the established FST2 showed that the two methods are comparable, if the 0.4 log unit (or 4 dB) bias on the D-FST is taken into account.

A possible reason for this bias could be the different flash durations used in the two methods. The FST2 uses 200-ms flashes for stimulation, above the normal temporal integration of human rod photoreceptors of 100 ms [14], while the default setting of 4 ms for the D-FST falls below rod integration time. Converting threshold units to log cd/m² biases the thresholds against the D-FST, since approximately half of the quanta in the 200-ms flash from the FST2 fall outside the rod integration time. However, for the conversion, the full 200 ms are assumed, leading to a lower threshold in cd/m² than warranted. This would account for a 0.3 log unit difference and would mean that Bloch's law does not apply to the 200-ms stimulus of the FST2.

Another possible reason for the bias is that the different psychometric procedures could influence the result. The FST2 uses a staircase method with 4-dB steps in the initial phase and 2-dB steps in the reversal phase. If a subject makes a mistake during the reversal phase, it can have a large influence on the final threshold. The D-FST on the other hand finds the approximate threshold, derives an intensity-response function and accounts for false positives and false negatives through the Weibull function. This could lead to a slightly more conservative measure of the threshold than on the FST2 and could explain the slight bias.

A third possible explanation for the bias could be that the 4-ms stimulus of the D-FST is in actuality not 4 ms long. Considering a possible ramp-up and ramp-down time for the electrical charge and discharge, the actual flash could be shorter and thus the conversion to cd/m² assuming 4 ms would lead to an overestimation of the threshold on the D-FST.

The FST2 has been used successfully to define residual vision in subjects with LCA [15] and has been validated as an outcome measure in treatment trials for RDD [9]. However, since the test is not widely available, it would be difficult to use as an outcome measure for a multi-center trial. Having established that the D-FST is an accurate and reliable psychophysical method, comparable to the FST2, while being incorporated in a widely used commercially available system (Espion E², Diagnosys LLC) and made more customizable for the tester, the D-FST represents a viable alternative for future treatment trials involving severely affected subjects. Finally, a recent update to the software further extends the stimulus range of the D-FST by incorporating intense flashes (up to 4.8 log cd/m²) generated by the xenon tubes.

Acknowledgments

We would like to thank Hemaxi Patel for her help with testing and both the patients with retinal degenerative diseases and normal subjects for participating in this study. This study was supported by funds from the Foundation Fighting Blindness.

References

1. Inoue Y, Iriyama A, Ueno S, Takahashi H, Kondo M, Tamaki Y, Araie M, Yanagi Y. Subretinal transplantation of bone marrow mesenchymal stem cells delays retinal degeneration in the RCS rat model of retinal degeneration. *Exp Eye Res.* 2007; 85(2):234–241. [PubMed: 17570362]

2. Maguire AM, Simonelli F, Pierce EA, Pugh EN Jr, Mingozzi F, Bennicelli J, Banfi S, Marshall KA, Testa F, Surace EM, Rossi S, Lyubarsky A, Arruda VR, Konkle B, Stone E, Sun J, Jacobs J, Dell'Osso L, Hertle R, Ma JX, Redmond TM, Zhu X, Hauck B, Zeleniaia O, Shindler KS, Maguire MG, Wright JF, Volpe NJ, McDonnell JW, Auricchio A, High KA, Bennett J. Safety and efficacy of gene transfer for Leber's congenital amaurosis. *N Engl J Med*. 2008; 358(21):2240–2248. [PubMed: 18441370]
3. Acland GM, Aguirre GD, Ray J, Zhang Q, Aleman TS, Cideciyan AV, Pearce-Kelling SE, Anand V, Zeng Y, Maguire AM, Jacobson SG, Hauswirth WW, Bennett J. Gene therapy restores vision in a canine model of childhood blindness. *Nat Genet*. 2001; 28(1):92–95. [PubMed: 11326284]
4. Bainbridge JW, Smith AJ, Barker SS, Robbie S, Henderson R, Balaggan K, Viswanathan A, Holder GE, Stockman A, Tyler N, Petersen-Jones S, Bhattacharya SS, Thrasher AJ, Fitzke FW, Carter BJ, Rubin GS, Moore AT, Ali RR. Effect of gene therapy on visual function in Leber's congenital amaurosis. *N Engl J Med*. 2008; 358(21):2231–2239. [PubMed: 18441371]
5. Hauswirth WW, Aleman TS, Kaushal S, Cideciyan AV, Schwartz SB, Wang L, Conlon TJ, Boye SL, Flotte TR, Byrne BJ, Jacobson SG. Treatment of leber congenital amaurosis due to RPE65 mutations by ocular sub-retinal injection of adeno-associated virus gene vector: short-term results of a phase I trial. *Hum Gene Ther*. 2008; 19(10):979–990. [PubMed: 18774912]
6. Rizzo JF 3rd, Wyatt J, Loewenstein J, Kelly S, Shire D. Methods and perceptual thresholds for short-term electrical stimulation of human retina with microelectrode arrays. *Invest Ophthalmol Vis Sci*. 2003; 44(12):5355–5361. [PubMed: 14638738]
7. Caspi A, Dorn JD, McClure KH, Humayun MS, Greenberg RJ, McMahon MJ. Feasibility study of a retinal prosthesis: spatial vision with a 16-electrode implant. *Arch Ophthalmol*. 2009; 127(4):398–401. [PubMed: 19365014]
8. Roman AJ, Schwartz SB, Aleman TS, Cideciyan AV, Chico JD, Windsor EA, Gardner LM, Ying GS, Smilko EE, Maguire MG, Jacobson SG. Quantifying rod photoreceptor-mediated vision in retinal degenerations: dark-adapted thresholds as outcome measures. *Exp Eye Res*. 2005; 80(2):259–272. [PubMed: 15670804]
9. 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. [PubMed: 17664667]
10. Swanson WH, Birch EE. Extracting thresholds from noisy psychophysical data. *Percept Psychophys*. 1992; 51(5):409–422. [PubMed: 1594431]
11. Bloch AM. Experiences sur la vision. *CRSM Soc Biol* 8/TII. 1885; 28:493–495.
12. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*. 1986; 1(8476):307–310. [PubMed: 2868172]
13. Peters AYL, Locke KG, Birch DG. Comparison of the Goldmann–Weekers dark adaptometer and LKC Technologies Scotopic Sensitivity tester-1. *Doc Ophthalmol*. 2000; 101(1):1–9. [PubMed: 11128963]
14. Barlow HB. Temporal and spatial summation in human vision at different background intensities. *J Physiol*. 1958; 141(2):337–350. [PubMed: 13539843]
15. Jacobson SG, Aleman TS, Cideciyan AV, Roman AJ, Sumaroka A, Windsor EAM, Schwartz SB, Heon E, Stone E. Defining the residual vision in Leber congenital amaurosis caused by RPE65 mutations. *Invest Ophthalmol Vis Sci*. 2009; 50:2368–2375. [PubMed: 19117922]

Fig. 1.

Representative D-FST functions in patients. The *solid black bar* denotes the normal range, the *white square* within the *bar* the normal mean. The *solid curve* is the Weibull function. Data-points are shown as *solid triangles*, which are partially obscured at 0%. **a** D-FST of a 5-year-old female with LCA. Visual acuity was LP, and both rod and cone ERGs were non-detectable (<2.0 and <0.1 μV , respectively). Her D-FST threshold of -22.5 dB (*solid black rectangle*) corresponds to -0.85 $\log \text{cd/m}^2$. **b** Example of a D-FST from a second patient where the floor did not reach 0% probability due to patient errors (indicated in *box*). The Weibull function adjusts the floor, ceiling and threshold accordingly

Fig. 2.

Box plots of the thresholds for subgroups tested on both systems. Groups are separated by their diagnosis and by system. *Boxes* show 25–75% range of values, *whiskers* denote full range. *adRP* autosomal dominant retinitis pigmentosa, *arRP* autosomal recessive RP, *isoRP* isolate RP, *LCA* Leber congenital amaurosis, *Usher* Usher Syndrome, *CRD* cone-rod dystrophy

Fig. 3.

a Scatterplot of thresholds for all patients tested on the FST2 (y -axis) plotted against the threshold for the same patient on the D-FST (x -axis). Results from the two tests show good agreement ($r = 0.91$). **b** Bland–Altman analysis of thresholds from the same patient on both tests. There is a slight bias, with thresholds on the D-FST being 0.4 log higher than on the FST2. *Top and bottom dashed lines* mark the 95% confidence range

Fig. 4. Bland–Altman analysis for the test/retest variability of the D-FST. The difference in threshold between test and retest within a session (y -axis) is plotted against the mean threshold for all sessions (x -axis). The 95% confidence range (*dashed lines*) is ± 0.3 log cd/m²

Fig. 5. Longitudinal D-FST data from a male (#8093) and two females (#8095, #8218) with isolate RP. Their visual acuity was limited to LP only, and both rod and cone full-field ERGs were non-detectable (<2.0 and <0.1 μV , respectively). The *star* denotes an episode of viral conjunctivitis. There was an increase in threshold of 0.6 log unit over 18 months for #8093, while visual acuity measurements remained at LP. Patients #8095 and #8218 remained stable within the 0.3 log confidence range

Table 1

Subjects are separated by diagnosis

Diagnosis	Subjects (n)	D-FST eyes (n)	FST2 eyes (n)	Mean age (years)	Median VA (logMAR)	Mean rod ERG (Mv)
Normal	7	7	7	39.3 ± 13.8	0 ± 0.1	106.4 ± 18.4
Autosomal dominant retinitis pigmentosa	5	8	4	61.8 ± 12.8	0.9 ± 0.8	2.6 ± 5.8
Autosomal recessive retinitis pigmentosa	7	10	4	50.7 ± 12.2	1.1 ± 0.8	0
Isolate retinitis pigmentosa	9	10	4	56.4 ± 17.9	1.4 ± 0.8	0
X-linked retinitis pigmentosa	3	6	0	64.0 ± 4.0	1.4 ± 0.7	0
Leber congenital amaurosis	6	7	4	18.9 ± 9.7	1.6 ± 0.9	0
Usher syndrome	6	6	2	41.5 ± 9.9	0.6 ± 0.6	20.1 ± 49.2
Bardet-Biedl syndrome	2	2	0	43.5 ± 3.5	1.9 ± 0.2	0
Cone-rod dystrophy	4	4	3	47.3 ± 31.4	1.6 ± 0.4	15.6 ± 24.3

Number of patients per condition is listed with number of eyes tested per method. Mean age, median visual acuity and mean dark-adapted rod ERG amplitude are listed with standard deviations. Affected subjects with measurable ERGs met the visual field criteria. Values assigned for acuities unmeasurable by ETDRS vision chart were as follows: *CF* 1.7 logMAR, *LP* 2.0 logMAR and *NLP* 3.0 logMAR