Details of testing methods

Full Field Stimulus Light Threshold

An Espion ColorDome™ Ganzfeld and the E2 console were used for stimulus presentation. A similar test has also been used to quantify visual sensitivity in gene therapy studies in Leber congenital amaurosis (LCA) clinical trials.[15, 16, 28] In a comparative investigation, both of these tests have been shown to output similar light thresholds.[29] Threshold was defined as the stimulus intensity at which the light was detected 50% of the time. Quantified light thresholds are reported only for subjects at sites equipped with the Espion system at the time of testing and whose light thresholds were measurable (pre-implantation: n = 8 implanted eyes, n = 8 fellow eyes; post-implant: n = 15 implanted eyes, n = 17 fellow eyes. Note: for two subjects, fellow eyes had no measurable threshold prior to implantation but did have quantifiable thresholds post-implantation. This is likely due to a change in the available maximum luminance rather than an actual change in light threshold and thus these eyes were excluded from statistical analysis). Two statistical analyses were performed on the FST data: (1) To determine whether light thresholds on each eye were correlated (across all subjects), the correlation coefficient (Pearson's r) was calculated and tested for significance (two-tailed); and (2) Pre- and post-implant (latest follow-up time point available) light thresholds in each eye were tested for significant difference over all subjects with a Student's two-tailed, paired t-test.

Perceptual Thresholds for Electrical Stimulation

The perceptual threshold to electrical stimulation of each individual electrode was measured using custom-developed software. Each stimulus presentation consisted of a 250 ms duration train of cathodic-first, biphasic pulses delivered at a frequency of 20 Hz (i.e., five pulses per train); the pulse width of each phase of individual pulse within a train was 0.45 ms and there was no interphase interval. This measurement was repeated at each follow-up time point, starting at 1 month after implantation.

Square Localization

Subjects viewed a 19” diagonal (48.2 cm) touch screen monitor (Elo TouchSystems AccuTouch; Tyco Electronics, Berwyn, PA) from 12’’ (30.5 cm). On each trial, a white 2.75 (7 cm) square was displayed at a random location on a black background and the subject was instructed to find and touch the square.

A test consisted of 40 trials with automated verbal feedback to indicate 1) whether the subjects touched the square and if not, 2) whether they were close, (within 100 pixels of a border of the square), and 3) the general direction (e.g., up and to the left) of the square relative to their touch response. For each trial, the distance between the center of the target square and the subject’s response was calculated; smaller values therefore indicate more accurate results (because the subject’s response was close to the target). The test was repeated with the system ON and OFF for each subject. Subjects were tested binocularly during all runs (system ON and OFF).

Direction of Motion

Subjects viewed the touch screen monitor as in the Square Localization task, and were instructed to maintain head (camera) fixation on the center of the screen during the test. On each trial, after an audio prompt, a white 1.4’’ (3.6 cm) line (oriented orthogonally to its path) swept across the monitor at a randomly chosen direction from 0 to 360°. After the stimulus presentation, subjects indicated the direction of motion they perceived by sweeping their finger across the touch screen. The speed of the motion was fixed throughout the test, but varied across subjects (the optimum speed for each subject was selected from three options: 1 second, 2 seconds, or 8 seconds to cross the screen).

A test consisted of 80 trials with automated verbal feedback, which indicated to the subject whether they detected the correct direction, or, if not, which of eight principal directions most closely matched the actual direction. For each trial, the difference between the stimulus angle and the response angle

was calculated; smaller values therefore indicate a more accurate result (because the response angle was close to the stimulus angle). The test was repeated with the system ON and OFF for each subject. Subjects were tested binocularly during all runs (system ON and OFF).

Grating Visual Acuity

The grating visual acuity test was developed to measure subjects' visual acuity; it was based on the principles of standard optometric acuity tests, but modified for vision lower than what is measurable with an ETDRS chart. Because there is no zoom or magnification in the conversion of the video signal to electrical stimulation in the implant, the visual acuities measured in prosthesis subjects are comparable to standard visual acuity measurements.

The grating visual acuity test was a four-alternative forced-choice test that presented gratings of black and white bars on a 20" (50.8 cm) flat screen monitor (Dell UltraSharp; Dell, Round Rock, TX) in one of four orientations (horizontal, vertical, and diagonal up to the left or right). The bars were presented for five seconds; subjects were required to provide a response to each trial, even if it was a guess. The widths of the bars were varied in 0.1 log unit steps according to an adaptive algorithm based on maximum likelihood in order to estimate the visual acuity threshold. Subjects completed the test with the non-tested eye patched (i.e., monocularly). They completed the test 3 times: (1) Implanted eye, system ON; (2) Implanted eye, system OFF; and (3) Non-implanted eye, system OFF.

Ninety-five (95) trials were presented in each condition, and based on these results a visual acuity score was given in units of the log of the minimum angle of resolution (log MAR), ranging from 2.9 to 1.6 log MAR; a confidence interval for that score was also determined. Visual acuity results were considered valid if they had a 95% confidence interval (computed using the subject's maximum performance) that was fully contained within the tested scale (2.9 – 1.6 logMAR). If the adaptive algorithm determined

that the subject's performance was no better than chance, the test was terminated early and the subject's acuity was scored as >2.9 log MAR.

Orientation and Mobility

These tests are performed with both eyes open to best represent real-world conditions. (Of course, since the device is implanted in only one eye, system ON viewing is not binocular in the typical sense.)

Door Test: In this test, subjects were asked to walk across a room and touch a piece of black or white felt (this represented a door) affixed to a light or dark wall (respectively) at the far end of the room. Subjects were placed in one of 3 different start positions (directly in front of the door or slightly to the left or right of the door) and they completed two trials from each start position, for a total of six trials per test. Midway through the study it appeared that some subjects, possibly due to increasing familiarity with the test environment, were able to successfully walk to the door with no visual input. Therefore, for the remainder of the study, this possibility was eliminated by moving the location of the door for each trial, and placing the subjects in a single start position. A success was scored if the subject was touching any part of the door at the end of the trial.

Line Test: In this test, subjects were in a room where black tile mats were placed on the floor to provide a consistent walking surface. The mats were painted with a white line and subjects were asked to follow the line to its end. Subjects were placed in one of 3 different start positions (straight ahead, facing slightly to the left or slightly to the right) and they completed two trials in each condition (for a total of 6 trials per test). Partway through the study, to reduce the likelihood that subjects would accidentally follow the straight line to its end, this test was also made more challenging by placing the subject at the starting point of the line, facing in the correct direction, but using three line configurations: straight, or with a 90° turn midway, either to the right or to the left, rather than completing all trials with a straight

line. A success was scored if the subject's final position before stepping off the tiled area was on the tile where the line ends.

At each follow-up time point, subjects performed the door and line tests with the Argus II System ON and OFF. RM ANOVA means of the percent-success rate were computed using a repeated measure analysis of variance (RM ANOVA). The observed means are known to be potentially biased when data are missing and there are (non-zero) correlations among patients from one time to the next. In contrast, means estimated using the repeated measure model are maximum likelihood (ML) estimates. ML estimates are known to be unbiased in most situations.[30]

The use of RM ANOVA requires that, controlling for time, the data follow a normal (Gaussian) distribution; the validity of this assumption was assessed with the Shapiro-Wilk statistic. The appropriateness of the RM ANOVA was further assessed by calculating the correlations across time; correlations near 1.0 show that a repeated measure model is helpful as there are non-zero correlations among patients across time.

RM ANOVA mean percent differences were tested against zero at each follow up visit; mean percent differences significantly lower than zero indicate that the outcome was significantly better with the system ON compared to OFF.