The Argus II epiretinal prosthesis system allows letter and word reading and long-term function in patients with profound vision loss

Lydon da Cruz,1 Brian F Coley,2 Jessy Dorn,3 Francesco Merlini,2 Eugene Filley,4 Punita Christopher,3 Fred K Chen,1,5 Varalakshmi Wuyyuru,3 Jose Sahel,6 Paulo Stanga,7 Mark Humayun,3,8 Robert J Greenberg,3 Gislin Dagnelie,9 for the Argus II Study Group

ABSTRACT
Background Retinal prosthesis systems (RPS) are a novel treatment for profound vision loss in outer retinal dystrophies. Ideal prostheses would offer stable, long-term retinal stimulation and reproducible spatial resolution in a portable form appropriate for daily life.

Methods We report a prospective, internally controlled, multicentre trial of the Argus II system. Twenty-eight subjects with light perception vision received a retinal implant. Controlled, closed-group, forced-choice letter identification, and open-choice two-, three- and four-letter word identification tests were carried out.

Results The mean ± SD percentage correct letter identification for 21 subjects tested were: letters L, T, E, J, F, H, I, U, 72.3±24.6% system on and 17.7±12.9% system off; letters A, Z, Q, V, N, W, O, C, D, M, 55.0±27.4% system on and 11.8±10.7% system off, and letters K, G, X, B, Y, S, P, 51.7±28.9% system on and 15.3±7.4% system off. (p<0.001 for all groups).

A subgroup of six subjects was able to consistently read letters of reduced size, the smallest measuring 0.9 cm (1.7°) at 30 cm, and four subjects correctly identify unrehearsed two-, three- and four-letter words. Average implant duration was 19.9 months.

Conclusions Multiple blind subjects fitted with the Argus II system consistently identified letters and words using the device, indicating reproducible spatial resolution. This, in combination with stable, long-term function, represents significant progress in the evolution of artificial sight.

Retinal prosthesis systems (RPS) are a new treatment for profound vision loss in retinitis pigmentosa.1 2 Ideal RPSs would provide stable, long-term retinal stimulation inducing visual percepts that allow form recognition while being safe, practical and portable, so it can be used in the setting of the subject’s daily life.3 4 Many of these features have been demonstrated in small studies, in differing combinations to date.5–11

Outcome measures of RPSs in blind subjects are being sought, with grating visual acuity, motion detection, tests of orientation and mobility, tasks of daily living, as well as quality-of-life outcomes being investigated.12–15 We have chosen letter identification as an outcome measure for spatial resolution in this study as it is currently the most familiar method in conventional clinical trials.16–21 However, standard Early Treatment Diabetic Retinopathy Study (ETDRS) tests are not validated or appropriate for prosthetic retinal devices, and as such, we have used a specific methodology that allows the identification of letters to be assessed using such devices.

The Argus II retinal prosthesis system (Second Sight Medical Products, Sylmar, California, USA) is fully portable and consists of a video camera mounted on glasses, a video processing unit (VPU) and a 60-electrode, epiretinal stimulating array.10 13 The array stimulates residual inner retina directly, mainly ganglion and possibly bipolar cells,22–24 with the signal transmitted via the normal visual pathway to the visual cortex. The aim of the tests reported here was to determine whether a group of patients blinded by outer retinal dystrophies and fitted with the Argus II system, could identify letters and words on a computer screen as a demonstration of effective spatial resolution, in the context of a device that is stable, portable and capable of long-term function.

METHODS
The Argus II feasibility study is a prospective, multicentre, internally controlled trial registered at (http://clinicaltrials.gov/show/NCT00407602). The institutional review boards or ethics committees, and competent authorities of participating sites and countries, approved the protocol. The trial adhered to the tenets of the Declaration of Helsinki.

Subjects/device
Twenty-nine subjects with retinitis pigmentosa, and one with choroideremia with vision reduced to light perception or worse, defined as unable to see gratings of 2.9 logMAR (20/16 000)20 21 in either eye, were enrolled in the trial. Two patients are not reported as one had the device explanted due to conjunctival erosion, the other had a retinal detachment. The remaining 28 subjects had functioning electrodes and reportable visual phosphenes at the end of this study, and were documented for long-term function. Twenty-one subjects took part in the letter-reading study. The nine subjects not included are one with silicone oil in situ due to hypotony, one awaiting replacement of their retinal tack, two who declined to participate due to the length of time taken to identify letters during training, three who were excluded due to inconsistencies in executing the test protocol and the two described above. For
each subject, age at diagnosis, age at implantation and the number of months the device had been implanted by the end of the study were recorded. The device consisting of video camera glasses, a VPU and a wireless-activated 60-electrode stimulating array has been described previously.10 13

Letters, monitor and lighting
Letters and words were white on a black background on an external LCD screen with resolution set to 1024x768 pixels (MatLab, Mathworks, Natick, Massachusetts, USA). Optotypes in True Type Century Gothic (CG) font were created in PowerPoint (Microsoft, Redmond, Washington, USA) and viewed at 30 cm. In Test I, the letter size was 41.27°. In Test II, the size was reduced from 41.27° in 0.1 log unit steps to 0.95°. In Test III, the sizes were 33.39° for the two-letter words, 26.81° for the three-letters words and 21.33° for the four-letter words. Testing was carried out in a darkened room.

Training for each test
Standardised training for Test I consisted of showing each subject each of the letters once with the RPS device on and telling them the name of the letter displayed while the patient scanned the screen. There was no further practice carried out before the formal testing. There was no training for Test II, the letter size reduction, or Test III, the word reading.

Test I: letter identification
All subjects were asked to identify all 26 letters. Letters were separated into three groups with differing typographical complexity. Group A only horizontal and vertical components (eg H, I) Group B oblique components involving the full height of the letter (eg, A M W), or letter with a minor variation on a circle (eg, O D C), and Group C letters with an oblique or curved element involving half the letter height (eg, K R) These were presented in a forced-choice, closed-set test in the order A then B then C. Researchers presented letters in random order, one at a time, four times each, both with the system on and off. Subjects who found the earlier groups difficult declined to continue the study for the later groups of letters. Twenty-one subjects completed Group A letters T, E, J, F, H, I, U; nineteen completed Group B letters A, Z, Q, V, N, W, O, C, D, M, and twenty Group C letters K, R, G, X, B, Y, S, P. The time taken to identify each letter was recorded.

Test II: letter size reduction
A subset of six subjects who in Test I correctly identified 50% or more of the letters in each group within 60 s, also completed Test II. Researchers asked subjects to identify the letters C, D, H, K, N, O, R, S, V, Z in a forced-choice, closed-set test; subjects were allowed to guess. Letters were presented individually, in random order, with five letters of each size constituting a ‘line’, and reducing size in 0.1 log unit steps. A time limit of 60 s per letter was set. The stopping point was five errors in any line, or when the subject felt that they were unable to guess the letter. Researchers carried out test and control runs twice, on separate days.

The total number of letters correctly identified, and height of the smallest correct letter (minimal letter size (MLS)) were recorded. The equivalent number of correct lines was calculated by dividing the number of correct letters by 5. The letter size of the smallest line using this calculation was recorded as the optimal letter size (OLS).

Test III: word recognition
Four subjects who scored more than 10 letters correct in Test II completed Test III. Common two-, three- and four-letter words in English or French, from word frequency tables, were presented in the subject’s preferred language. Letter spacing was standard for the font size with no additional spacing created between the letters. A time limit of 120, 180 and 240 s for two-, three- and four-letter words was set with a warning 10 s before the end.

Test and control conditions
For Test II letters and Test III words, the five tested conditions were: system off and both eyes unpatched; system on in ‘scrambled’ mode with eyes patched and unpatched; and system in standard mode with eyes patched and unpatched. The scrambled mode13 consisted of the system being on but the spatial mapping of the image onto the electrodes randomised so the spatial correlation between camera image and retinal stimulation was lost (figure 1D).

Statistical analysis
Researchers used the Wilcoxon signed rank test to examine the significance of the difference between conditions for each of the groups tested, and standard multiple regression to determine how much variance in the percentage of correct letter identification could be explained by three independent variables: age at implantation, age at diagnosis and number of electrodes activated. For Test II, a statistical analysis was done by using a one-way analysis of variance (ANOVA) model to determine whether the number of correct answers was significantly influenced by the system condition. Second, another one-way ANOVA analysis was performed to determine whether the number of correct answers (with the system ON and standard mode) was significantly influenced by the different patching conditions.

RESULTS
Subjects and implant
Gender, age at diagnosis and at implantation, duration of implantation and number of electrodes included in the video configuration file at the time these tests were conducted are recorded in table 1. Average implantation duration was 19.9 months (range 8.6–34.8) at test completion, with all devices demonstrating functioning electrodes throughout testing.

Test I
Mean percentage correct±SD for subjects tested were: Group A 72.3±24.6% system on and 17.7±12.9% system off; Group B 55.0±27.4% system on and 11.8±10.7% system off and Group C 51.7±28.9% system on and 15.3±7.4% system off. (p<0.001, device on versus off for all groups—Wilcoxon signed rank test). The median and quartile ranges for outcomes are shown in figure 2. The average times (in seconds) for correctly identified letters (system on) were 47.7, 68.6 and 63.9 for Groups A, B and C, respectively.

Test II
The MLS and OLS in cm (degrees) viewed at 30 cm for subject 51-009 are 0.9 (1.77) and 5.6 (10.6°); 52-001, 1.8 (3.4°) and 3.6 (6.8°); 61-005, 1.1 (2.1°) and 2.3 (4.3°); 61-003, 3.6 (6.8°) and 14.3 (27.1°); 12-005, 3.6 (6.8°) and 14.3 (27.1°); and 52-003, 18.0 (32.4°) and 22.6 (42.9°). The average number of
letters correctly identified for each test modality is summarised in figure 3.

Test III
The tabulated results of the number correct in word recognition for two-, three- and four-letter words for each of the test situations is reported in table 2. The results show there is a clear improvement for all four subjects with the device on (both patched and unpatched).

Multiple regressions (Test I)
After adjusting for the small sample size, the three variables, age at implantation, age at diagnosis, and number of electrodes activated jointly, explained 19% (ns), 27% (ns) and 39% (p<0.01) of the variances for correct letter identification in Test I, Groups A, B and C letters, respectively. For all three groups of letters, number of electrodes and age at the time of implantation made the largest unique contributions to the percentages above. Only age at implantation correlated singly with statistical significance and explained 30% (ns), 36% (p<0.01) and 28% (p<0.01) of the total variances for Groups A, B and C letters, respectively.

DISCUSSION
The research team have investigated the Argus II system both in terms of spatial resolution, using letter identification, and long-term stable function. In Test I on average, across all the subjects, letters could be clearly identified correctly between 51.7% (Group C) and 72.3% (Group A) of the time with the system on while performing at chance level with the device off. The median values for correct letters, system on, in figure 2 of 50% or greater for groups A, B and C indicate that at least half the group was functioning at this level (figure 2).

It is possible that a component of letter identification involves head scanning with light detection and inference of the letter. This may partially account for the greater number correct for Group A letters when the Test I letters were stratified into three groups of increasing typographic complexity. However, experimenters observed that many high-performing subjects did not scan across the letter but followed the form of the letter as if tracing it, suggesting that they were able to see the straight and curved strokes that make the form of the letter. Furthermore, failure to identify letters in Test II with the scrambled mode (not better than device off) showed that letter identification is not primarily dependent on head scanning, light detection and inference, but uses the spatial information contained in the percept. The two outlier patients (figure 2 device off) may have used residual light perception to construct a sense of the letter to aid guessing in the less complex letters in Groups A and B. However, in both subjects, the letter-reading performance was both better and faster with the device on.

Test II examined six high-functioning subjects in terms of the limit of letter size identification. The study team used the framework of the ETDRS letter set and layout, with five letters per ‘line’ and 0.1 log unit steps. CG font was chosen to maximise the chance of recognition with an RPS as the narrow lines of the font maximised the space between linear elements and presented a low mean illumination. Unlike the standard ETDRS test, the letters were presented one at a time in white on a black background.

The best subjects were able to consistently identify letters from 2.3 cm, (4.3°) to 22.6 cm, (42.9°) in height (OLS). The smallest letters identified ranged from 0.9 cm (1.7°) to 18 cm (34.2°). As the electrode spacing in the array is 1.8°, this shows that, in at least some subjects, the system is working at a level approximating predicted limit of spatial resolution. This was also shown previously for the Argus I device.25 As expected, there were significant errors at the smaller sizes. The scrambled mode control confirmed that letter recognition was based on true spatial resolution and not light detection. Patching confirmed that the outcomes were not achieved due to residual
native light perception or a placebo effect. We feel that the OLS gives a better functional indication of the size where letter identification was consistent. We used letter height as the outcome, as CG letters do not have the equivalent defining elements within each character as Sloan letters in the ETDRS set. As such, the outcome is a descriptive indication of the functional level of letter identification possible rather than a formal visual acuity measure.

Four subjects from the same high-performing group could also read short words. This demonstrated the ability to define separate letters with minimal spacing and not to confuse multiple letters. Due to the small sample size, a statistical analysis was not performed. Nevertheless, these results support the proof of principle that this device allows simple word reading. True spatial resolution was indicated, as there was no consistent word reading with the scrambled control. Letter reading has been previously reported in a patient with a subretinal implant with resolution 'corresponding to a height of approximately 9° of visual angle'. The subject was described as being able to identify 16 different cut-out letters, and read five words correctly 'repeatedly on several days'. A few words were correctly identified when viewed with eyes unpatched and the device off or in scrambled mode. We feel that the residual light perception, with or without the Argus II acting as a light detector, may have led to a sense of the letter or word allowing enhanced guessing. With the device in scrambled mode and both eyes patched, the two correct identifications were likely lucky guesses.

Despite the good overall level of letter identification, the average time taken to correctly identify letters in Test I ranged across subjects from 6 s to 221 s. What differentiates high-functioning subjects from others is not clear at present. Multiple regression analysis suggests that age at implantation correlated to letter reading outcome with statistical significance. This may reflect the presence of healthier residual inner retina. Age at diagnosis and number of electrodes activated (see table 1) also correlated to better outcome with letter identification, though not statistically significantly, likely reflecting the small sample size. Genotype may play a part, but the number of cases in this study with a known genotype precludes investigation of this issue. Identifying these factors will be important in informing...
future case selection. Also, it is not immediately clear how the performance of the controlled tasks identified in this study will translate directly into useful function in daily life, and this is being studied further.

The Argus II RPS was stably fitted into the 28 subjects for an average time of 19.9 months (range 8.6–34.8) from implantation to the end of the tests in this study. Throughout the study, 28 out of 30 implants had functioning electrodes (as recorded in table 1), thus demonstrating the possibility for stable chronic stimulation. Of the other retinal prosthetic trials, the Optobionics chip was reported stable in the eye for up to 18 months, the Argus I to more than 5 years, and the subret implant study up to 5 weeks.

CONCLUSION
This study demonstrates that the Argus II RPS shows features required of a successful artificial retina system. In accordance with this, it became the first retinal prostheses to receive regulatory approval for implantation in Europe outside of a trial, in April 2011. This paper highlights a large number of subjects that are able to identify letters, and a group identifying words correctly by the four subjects in five conditions. The test conditions are indicated by the column labels.

Table 2  Word reading outcome table

<table>
<thead>
<tr>
<th>Subject</th>
<th>System off</th>
<th>System on scrambled mode</th>
<th>System on standard mode</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Unpatched</td>
<td>Patched</td>
<td>Unpatched</td>
</tr>
<tr>
<td>Two-letter words (10 words per trial)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>51-009</td>
<td>2</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>52-001</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>61-003</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>61-005</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Three-letter words (10 words per trial)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>51-009</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>52-001</td>
<td>1</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>61-003</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>61-005</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Four-letter words (10 words per trial)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>51-009</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>52-001</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>61-003</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>61-005</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Table summarising the number (out of 10) of two-, three- and four-letter words read correctly by the four subjects in five conditions. The test conditions are indicated by the column labels.

References