Reproducibility of the visual evoked potential using a light-emitting diode stimulator

NEVELL F SKUSE, DAVID BURKE, BRIAN McKEON

From the Unit of Clinical Neurophysiology, Department of Neurology, The Prince Henry Hospital and School of Medicine, University of New South Wales, Sydney, Australia

SUMMARY The intersubject and intrasubject reproducibility of the VEP was evaluated using two different methods of producing pattern reversal—a mirror/projector system and a light-emitting diode (LED) system. Intersubject reproducibility was determined in 100 normal subjects (50 males, 50 females). Ten subjects were studied on ten different occasions over 11 months to establish intrasubject reproducibility. The two methods gave comparable results. It is concluded that the mirror/projector system has no advantages over the LED system, which is cheaper, more robust and more convenient. Most of the intersubject variability was found to be due to subject variables and it is concluded that technical improvements are unlikely to reduce further the variability of the test. A variability in the latency of P100 of up to 7.7 ms was recorded in the serial studies on the ten subjects, indicating that in longitudinal studies on patients, changes in latency must exceed 9–10 ms to be significant. With both stimulus systems, there were significant sex-related differences in latency (P50 and P100 but not N70) and amplitude (N70–P100). The extent of the difference was such that the upper limit of normal latency for P100 (2.5–3 SD from the mean) was 4.2–4.7 ms longer for males than females—a value which exceeds 1 SD. Separate control values for males and females are advisable whichever method is used to produce pattern reversal.

Since the studies of Halliday and colleagues1,2 the assessment of visual pathways using a reversing check pattern has become an almost obligatory component of the diagnostic work-up for patients with suspected multiple sclerosis. Many commercial firms now offer and many laboratories, particularly in the USA, have adopted a television system as visual stimulator instead of the original mirror/projector system used by Halliday’s group. However when tested on the same subjects the variability of the visual evoked potential (VEP) produced by the television system is greater than that produced by the mirror/projector system,3 so that this preference could compromise the diagnostic efficacy of the test. Light-emitting diodes (LEDs) have potential advantages over both a television screen and the original mirror/projector system. A visual pattern stimulator using round LEDs4 has produced results comparable to those of the mirror/projector system in one small series.5 However, if the stimulator was constructed using square LEDs,6 a more equitable comparison with the mirror/projector system would be possible. Such a stimulator has produced results apparently as satisfactory as those with the TV stimulator,7 although a detailed comparison was not undertaken.

If the choice of a pattern stimulator is determined by convenience and utility rather than proven low intersubject and intrasubject variability, the diagnostic precision of the test suffers. At the present time, there are no data indicating that other pattern stimulators can reproduce the low intersubject variance of the original mirror/projector system. Such data can be obtained only by comparing different stimulators in the same subjects under identical test conditions. Intrasubject variability is as critical as intersubject variability, particularly now that longitudinal studies are being performed on patients. However, there are few details on the reproducibility of results in normal subjects in the neurological literature, apart from the report of Stockard and colleagues7 on 10 normal subjects tested on two...
occasions and five on three occasions.

With any diagnostic investigation there are technical and subject contributions to the variability of the results obtained in a normal population. That gender contributes significantly to the variability of the VEP to pattern-reversal in a normal population is not accepted universally. It is said to do so by Stockard and colleagues,7 by Kjaer8 and by Halliday and colleagues,9 but this could not be confirmed in the recent report of Shearer and Dustman.10 Unfortunately, the report of Halliday et al.9 did not appear in conventional refereed literature and those of Stockard et al7 and Shearer and Dustman10 were published in a technical journal. Both could be lost to a potentially interested clinical audience.

The present study was undertaken, firstly, to compare the mirror/projector system with a LED system in normal subjects, secondly, to check the reproducibility of the findings with each stimulator in the same subjects on separate occasions and, thirdly, to determine whether there are significant differences between the sexes in the VEP.

Methods

Subjects The study population consisted of 50 male and 50 female subjects, aged 18-40 years. This age group was chosen because it corresponds to the age range for presentation of most patients with multiple sclerosis and because the latency of P100 does not change with age in this range.4,9 The subjects were healthy volunteers with no history of neurological or ophthalmological disorders. Fifteen wore spectacles during the tests to correct refractive errors. To determine reproducibility, five male and five female subjects were tested with both stimulators on 10 separate days over 11 months. The duration of each test lasted approximately one hour: no subject noted fatigue or discomfort in this time.

Stimulators The order of presentation of the two stimulators was randomised. At the end of the series, the mirror/projector system had been used first with 58 subjects, the LED stimulator first with 42 subjects. The mirror/projector system was a Devices Light Stimulator D110, as used by other authors.2,11 The overall screen presented to the subject contained 961 checks, with luminance of 560 cd/m² for the white squares and 24 cd/m² for the black squares (contrast ratio 23:1). The pattern reversal time was limited by the mechanics of the galvanometer, taking 8 ms to reach 90% of the full reversal distance. The LED stimulator and its driving circuitry were modified from those described by Epstein.6 The LED screen consisted of 144 Hewlett Packard HLMP 2655 red light emitting diodes in a 12 × 12 matrix, a larger screen containing more LEDs than previously described.6,7 The luminance of the red and black squares was 219 and 2×8 cd/m² respectively (contrast ratio 78). It has been shown that the red colour does not influence the amplitude of the visual evoked potential.12 and changes in contrast have little effect provided that the contrast ratio is over 20-40%.6 The pattern reversal time was faster than the response time of available light-dependent resistors (0-3 ms) and could not be measured accurately.

Experimental procedures The subjects were tested when lying supine on a comfortable bed with the head supported by pillows so that the neck flexed to 30°, a posture standard in our laboratory because severely incapacitated patients often experience difficulty in maintaining head and trunk stability when sitting. The screens of both stimulators were positioned so that each square subtended an angle of 1° at the retina, the overall screen size being 31° square and 12° square for the mirror/projector system and the LED system respectively. Stimulation of each eye was performed separately with vision of the other eye occluded by an eye shade. Pattern reversal occurred at 3/s for each system. The ambient room luminance was 100 cd/m² throughout the series.

The evoked potential was recorded using a pair of subcutaneous needle electrodes, the active at Oz and the reference at Fz. The amplifier band-width was 1-6 Hz-1-6 kHz. Two runs, each containing 256 responses, were summed using a fixed-programme averager with automatic overload rejection (Medelec DV62 and AX6). The sampling rate was 1-7 kHz. Each sweep and the updated average were monitored continuously on an oscilloscope.

Analysis Measurements were made following summation of the two averages. The latencies of P50 (when present), N70 and P100 were measured using a cursor. The amplitude of P100 was measured from the preceding negative peak (N70), as in fig 1. The inter-ocular latency difference in any one subject was calculated as longer latency minus shorter latency (regardless of actual size), and the symmetry of the amplitude of P100 in any one subject was...
Reproducibility of the visual evoked potential using a light-emitting diode stimulator

Table 1

<table>
<thead>
<tr>
<th>Subject</th>
<th>Left eye</th>
<th>M/P system</th>
<th>Significance*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>LED system</td>
<td>M/P system</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Latency (ms)</td>
<td>97.6 ± 3.76</td>
<td>98.5 ± 3.43</td>
</tr>
<tr>
<td></td>
<td>Amplitude (μV)</td>
<td>7.9 ± 3.24</td>
<td>11.4 ± 4.51</td>
</tr>
<tr>
<td>Right eye</td>
<td>Latency (ms)</td>
<td>97.1 ± 4.11</td>
<td>97.8 ± 3.58</td>
</tr>
<tr>
<td></td>
<td>Amplitude (μV)</td>
<td>8.2 ± 3.59</td>
<td>11.46 ± 4.44</td>
</tr>
<tr>
<td>Intercocular difference (ms)</td>
<td>1.4 ± 1.05</td>
<td>1.32 ± 0.98</td>
<td></td>
</tr>
<tr>
<td>Amplitude ratio</td>
<td>0.85 ± 0.13</td>
<td>0.88 ± 0.09</td>
<td></td>
</tr>
</tbody>
</table>

*Calculated using two-tailed t test for paired samples.

Table 2

<table>
<thead>
<tr>
<th>Subject</th>
<th>P100 (ms)</th>
<th>Interocular difference (ms)</th>
<th>Amplitude ratio</th>
<th>P100 (ms)</th>
<th>Interocular difference (ms)</th>
<th>Amplitude ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>96.9 ± 1.18</td>
<td>1.2 ± 0.85</td>
<td>0.8 ± 0.11</td>
<td>98.4 ± 1.99</td>
<td>1.2 ± 1.12</td>
<td>0.8 ± 0.1</td>
</tr>
<tr>
<td>2</td>
<td>97.1 ± 1.29</td>
<td>1.3 ± 0.91</td>
<td>0.9 ± 0.06</td>
<td>93.2 ± 1.38</td>
<td>0.7 ± 0.42</td>
<td>0.9 ± 0.05</td>
</tr>
<tr>
<td>3</td>
<td>97.2 ± 1.37</td>
<td>1.8 ± 1.76</td>
<td>0.7 ± 0.17</td>
<td>93.7 ± 1.87</td>
<td>2.2 ± 1.4</td>
<td>0.9 ± 0.08</td>
</tr>
<tr>
<td>4</td>
<td>100.8 ± 1.38</td>
<td>1.4 ± 0.53</td>
<td>0.9 ± 0.09</td>
<td>105.2 ± 1.48</td>
<td>1.7 ± 0.73</td>
<td>0.8 ± 0.06</td>
</tr>
<tr>
<td>5</td>
<td>92.3 ± 0.99</td>
<td>1.8 ± 1.07</td>
<td>0.95 ± 0.05</td>
<td>94.3 ± 1.36</td>
<td>1.4 ± 1.02</td>
<td>0.9 ± 0.06</td>
</tr>
<tr>
<td>6</td>
<td>92.7 ± 1.49</td>
<td>0.9 ± 0.52</td>
<td>0.9 ± 0.12</td>
<td>93.4 ± 1.77</td>
<td>0.8 ± 0.51</td>
<td>0.9 ± 0.11</td>
</tr>
<tr>
<td>7</td>
<td>91.4 ± 1.48</td>
<td>1.9 ± 1.13</td>
<td>0.7 ± 0.14</td>
<td>93.5 ± 0.98</td>
<td>1.3 ± 0.46</td>
<td>0.9 ± 0.1</td>
</tr>
<tr>
<td>8</td>
<td>101.8 ± 0.64</td>
<td>2.0 ± 0.61</td>
<td>0.9 ± 0.09</td>
<td>101.8 ± 0.97</td>
<td>1.2 ± 0.74</td>
<td>0.9 ± 0.08</td>
</tr>
<tr>
<td>9</td>
<td>90.3 ± 1.78</td>
<td>1.1 ± 0.91</td>
<td>0.8 ± 0.15</td>
<td>94.8 ± 2.07</td>
<td>1.8 ± 0.77</td>
<td>0.8 ± 0.09</td>
</tr>
<tr>
<td>10</td>
<td>100.4 ± 1.44</td>
<td>2.0 ± 1.66</td>
<td>0.9 ± 0.08</td>
<td>99.9 ± 1.2</td>
<td>1.1 ± 0.76</td>
<td>0.8 ± 0.09</td>
</tr>
</tbody>
</table>

expressed as the smaller amplitude divided by the larger (again regardless of the actual side).

In intersubject and intrasubject comparisons the latencies of the P100s for the two eyes of each subject were averaged; for example the means and standard deviations given for each subject in table 2 are based on ten sets of data, each the average P100 latency of the two eyes. In the tables, the values given represent mean ± standard deviation.

Results

Apart from latency and amplitude (see below), there were no consistent differences in the formats of the VEPs produced by the two stimulus systems—black-white pattern shift with the mirror/projector system, and black-red pattern alternation with the LED system. The primary aims of this study were (1) to compare the two methods of producing a visual evoked potential and (2) to investigate sex-related differences in the VEP.

The values for P100, the interocular difference, and the amplitude ratio are given in table 1 for each eye. If the values for the right and left eyes are averaged, the mean latencies of P100 were 97.4 ± 3.94 ms for the LED system and 98.2 ± 3.51 ms for the mirror/projector system. The latencies for P50 and N70 were also shorter using the LED system, by an average of 1.4 ms and 2.8 ms, respectively. The difference in the size of the discrepancy in the mean latencies of P50, N70 and P100 with the two systems was not due to greater variability of N70 (see table 3) and is consistent with the view that the successive waves of the VEP do not represent sequential activation of a chain of potential generators.13

The differences in the amplitude N70–P100 with the two systems were investigated further in a subgroup of 13 subjects. With the mirror/projector screen of normal size (31°, 961 checks) the amplitude of the potential was 9.0 ± 3.26 μV, significantly different (p < 0.02) from the amplitude produced by the LED system (mean 7.6 ± 2.3 μV). When the number of visible checks on the mirror/projector screen was reduced to 144 (the same number visible with the LED system) the amplitude of the response was reduced to 7.0 ± 2.45 μV. This value does not differ significantly from that produced by the LED system.

intersubject reproducibility

A difference in mean latency of P100 with the two systems does not indicate that one system is preferable to the other: the important parameter is the variance of the mean latency, not its absolute value. The smaller the standard deviation, the tighter the range of normality and the greater the potential diagnostic efficacy of the test. As shown in table 1, the variability of the mean latency of P100 was slightly smaller for the mirror/projector system. Analysis of variance using the “F” test revealed that the small differences in the standard deviations were not significant. The difference in the standard deviations (0.3–0.5 ms) amounts to only 0.5–1.0 averag-
ing bins, and it is concluded that in practice this difference is immaterial. There were also no significant differences in the variability of the other clinically useful parameters.

**INTRASUBJECT REPRODUCIBILITY**

Ten subjects, five males and five females, were each studied on ten separate days over a period of 11 months. Similar results were obtained with the two systems (table 2). The variability of P100 was smaller for seven subjects using the LED system, but the average standard deviations were close (1.38 ms for the LED system; 1.51 ms for the mirror/projector system), and the difference was not statistically significant (two sided t test). The range of variability of the ten tests in each subject was up to 7-7 ms with the LEDs and 6-9 ms with the mirror-projector system.

The degree of variability in the intrasubject data in table 2 allows an estimate of the contribution of subject variables and technical variables to the intersubject data. For example, with the LED stimulator, the mean standard deviation of 1.38 ms for the intrasubject data can be compared with the standard deviation of 3.94 ms for the intersubject data. The value 1.38 ms constitutes the variability due to the technical or non-subject influences. The subject contribution to the variability of the intersubject data can be calculated as follows:

Subject contribution = total variability minus non-subject contribution

\[ \text{subject contribution} = \sqrt{(3.94^2 - 1.38^2)} \]

\[ = 3.69 \text{ ms.} \]

Thus, of the standard deviation of 3.94 ms, approximately 3.69 ms can be attributed to subject variables. In practice therefore virtually all the measured variability in intersubject data is due to subject variables.

**SEX RELATED DIFFERENCES**

Similar differences between males and females were found in the pattern-reversal VEP using the LED stimulator and the mirror/projector stimulator. For brevity, only the data obtained with the LED stimulator will be presented in detail. The mean latency of P100 was significantly longer in males (table 3, fig 2). This difference was highly significant: for the LED stimulator—right eye p < 0.001; left eye, p < 0.001; and for the mirror/projector stimulator—right eye, p < 0.005; left eye, p < 0.001. The extent of this sex difference and the effect on the "limits of normality" can be illustrated by comparison of the separate male/female data with the pooled male/female data. The mean difference (2.5 ms for the right eye and 2.8 ms for the left eye) is over 60% of one standard deviation of the pooled normal range—4-11 ms and 3-76 ms respectively. Using the data from tables 1 and 3 for the right eye, the upper limits of normality (defined as 3 SD from the mean) would be 109-4 ms for males and females together, 111-7 ms for males, and 107-0 ms for females, so that a female patient with PT00 at 108 or 109 ms would be considered within normal limits if normality was defined without regard to sex. (If the upper limit of normality is set at 2.5 SD the relevant figures are 107-4 ms, 109-4 ms, and 105-2 ms respectively for the LEDs.) In either case the difference between the upper limits of normality for males and females (4-7 and 4-2 ms) exceeds 1 standard deviation.

The amplitude (N70–P100) also differed significantly between the sexes, with both stimulators. The model values were similar (7 μV), the higher mean amplitude of the VEP in females being due to a more skewed distribution, containing a greater number of large-amplitude potentials. In the low-amplitude range, the distributions for males and females were very similar.

Differences between the sexes tend to disappear in data on interocular symmetry. Nevertheless, the upper limits of the normal interocular latency difference (mean plus 3 SD) differ slightly for males and females: with the LEDs, 5-0 ms and 3-9 ms; with the mirror/projector, 4-7 ms and 3-6 ms respectively. Similarly, sex-related amplitude differences tend to
Reproducibility of the visual evoked potential using a light-emitting diode stimulator

When the amplitude ratio is used as a measure of amplitude symmetry. Again the lower limits of normality (means minus 3 SD) differ for males and females; with the LEDs, 0.41 and 0.54; with the mirror/projector, 0.59 and 0.65 respectively.

Significant sex-related differences were also found in the latency of P50 but not of N70 (table 3). The two stimulators yielded essentially identical findings. The absence of a sex difference for N70 cannot be attributed to a higher variability of the mean latency because the standard deviations for the three peaks are comparable (see table 3). In addition, there is not an orderly progression of increasing discrepancy between males and females for P50, N70 and P100. These findings support the view that the peaks P50, N70, P100 (and N135) do not result from sequential activation of a chain of neural generators.

Discussion

The present study has shown that there are no important differences between a LED stimulator and a mirror/projector stimulator for pattern-reversal VEPs, at least in a normal population. The results with the two systems are equally reliable and reproducible. In addition, the present study documents the intrasubject reproducibility of the VEP using two systems in a more detailed manner than has been previously published. A knowledge of the range of variability that can occur on different days in the VEP of a normal subject is critical if the results of longitudinal studies in patients are to be interpreted meaningfully. The present study has not established that the diagnostic yield of the LED stimulator in a patient population is the same as with the mirror-projector system. However, the comparable variances recorded with the two systems suggest that this is likely and, indeed, our subsequent experience indicates that this is so (unpublished observations).

The choice between a LED stimulator and the mirror/projector stimulator for routine use rests on factors other than their reliability and reproducibility. On these other grounds the LED system has clear advantages. The LED system is light, compact and cheap compared with the mirror/projector system. The LED matrix can be positioned easily to accommodate the patient whether he be reclining, supine or upright. The matrix is robust but the translucent screen of the mirror/projector system is easily broken. The response of the LED system is such that the pattern change is practically instantaneous but the mirror/projector system is limited by inertia of the moving parts. The LED system can be driven at very fast rates, well past the frequencies at which the human subject can perceive the reversal and, with it, conditioning-test paradigms and flicker-fusion can be studied. Finally, replacement of the globe, replacement of the screen, dust on the globe and decay of luminance as the lamp ages are all stimulus changes that are difficult to control or to correct for in the result of a test. Indeed, our control values for the mirror/projector system have changed.

Fig 2  Mean latencies of P100 for the left eye of males and females shown as a histogram on the left and a cumulative probability plot on the right.

<table>
<thead>
<tr>
<th>Male</th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>80</td>
</tr>
<tr>
<td>110</td>
<td>50</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>
in 7 years (during which time the lamp has been replaced four times and the screen twice) from a P100 of 94.8 ± 4.12 ms (n = 60) to the present value of 98.2 ± 3.51 ms.

The intrasubject comparisons document the reproducibility of the VEP in individual subjects. In this respect there were, again, no significant differences between the two systems. The range of variability was always less than 7.7 ms in the 10 tests on each of the 10 subjects. It can be concluded that providing testing conditions are identical and the patient co-operative, a change in the latency of P100 of 9–10 ms or more is likely to be significant. However, changes less than this can probably be ignored. The average standard deviation for P100 of 1.38 ms (LEDs) and 1.51 ms (mirror/projector) indicate that technical factors contribute little to the variability of P100. Of the 3–94 ms standard deviation for the LED system, 3.69 ms can be attributed to subject variables. This implies that further technical improvements are unlikely to reduce further the standard deviation for this form of testing. Attempts to improve the yield from VEP testing should be directed towards minimising subject variables, such as age and sex.

Sex differences in the latency of P100 make a significant contribution to the latency variability of a mixed population. Removal of this contribution by defining normal values for each sex significantly increases the power of the pattern-reversal VEP as a diagnostic tool. Similar results have now been found with three different stimulating systems—television,7 mirror/projector9 (and present study) and LED (present study)—on a total of 200 males and 204 females. The difference in the "upper limit of normal" of P100 for males and females exceeds one standard deviation of the mean.

The source of the differences in the latency of P100 between males and females has been attributed to differences in size and possibly differences in core temperature. Size, particularly skull thickness, is probably the major factor in the amplitude differences between the sexes. The female brain is smaller than the male brain, so the conduction pathway is likely to be shorter. However this is unlikely to be the sole factor in the shorter female P100 latency because a significant latency difference disappears after age 50, presumably due to menopausal changes. This change in the female P100 latency and the consequent disappearance of the sex-related latency difference may be of relevance when longitudinal studies are being undertaken.

The authors are grateful to Drs JW Lance and AK Lethlean for their support and comments on the manuscript.

References

17. Snyder EW, Dustman RE, Shearer DE. Pattern reversal evoked potential amplitudes: life span changes. Elec-
Reproducibility of the visual evoked potential using a light-emitting diode stimulator
