Reproducibility of the spike-triggered averaging technique of motor unit estimation (MUE) was assessed in biceps-brachialis muscle in 10 normal subjects and 15 subjects with amyotrophic lateral sclerosis (ALS). MUE was calculated by dividing the compound muscle action potential by the mean amplitude of 15 surface motor unit potentials (S-MUPs) of low recruitment threshold. Averaged MUE values in normal subjects were higher than in ALS subjects, with few values overlapping. Differences between test and retest MUE values were not significant for either subject group. The relative differences between test–retest values were 45.3% for normal subjects and 32.6% for ALS subjects. Correlation coefficients between test and retest values were low \( r = 0.07 \) for normal subjects when influential outlying points were removed, and higher \( r = 0.65 \) for ALS subjects when individuals with MUE values within the normal range were removed. The higher correlation of test–retest MUE values in ALS subjects compared to normal subjects may be due to a greater probability of resampling among the smaller number of motor units in ALS subjects. In summary, the reproducibility and technical aspects of the spike-triggered averaging technique are similar to those reported for other MUE techniques.

Key words: motor unit estimation • reproducibility • ALS

MOTOR UNIT ESTIMATION: REPRODUCIBILITY OF THE SPIKE-TRIGGERED AVERAGING TECHNIQUE IN NORMAL AND ALS SUBJECTS

MARK B. Bromberg, MD, PhD

An electrophysiologic method to estimate the number of motor units in a muscle which is convenient to perform, accurate, and reproducible would be a useful test. McComas and coworkers\(^\text{10}\) described a method of motor unit estimation (MUE) based on incremental motor nerve stimulation to arrive at the amplitude representative of single motor units. The mean representative single motor unit amplitude was divided into the maximal compound muscle action potential (CMAP) to obtain an estimate of the number of motor units making up the CMAP. Since then, several other methods have been proposed to determine the mean or representative amplitude of a single motor unit.\(^\text{5,6,8,12,18}\) The assumptions and limitations of each method have been reviewed.\(^\text{4,5,12,16,22,24}\)

The accuracy of the various methods cannot be assessed because there is no anatomic method to directly count the number of motor fibers innervating a muscle.\(^\text{22,24}\) A wide range of MUE values has been reported for the same muscle in normal subjects when different counting methods are used.\(^\text{1,5,18}\) An undetermined portion of the variability of MUE values may be due to imprecision of the test methods and not to biologic differences between subjects. Reproducibility of MUE methods becomes important when values are at the lower limit of the normal range because of the consequences of false positive and false negative diagnoses based on interpretation of the MUE values. Reproducibility is also important if MUE is to be used to follow the clinical course of patients who
have progressing or resolving denervating diseases, or to assess response in treatment trials.

There is little information available on reproducibility for the various MUE methods for normal subjects, and even less for subjects with denervating diseases.\textsuperscript{1,3,10,11,15-17,21} In this report, test–retest reproducibility of the spike-triggered averaging MUE technique\textsuperscript{6} is evaluated in normal subjects and in subjects with amyotrophic lateral sclerosis (ALS)-type motor neuron disease, who are expected to have low MUE values. The spike-triggered averaging MUE technique was selected because it can be performed easily on many EMG machines and requires no special or proprietary software for data analysis.

**METHODS**

Motor unit estimation reproducibility studies were carried out in the biceps-brachialis muscle group in 10 normal subjects whose mean age was 36 (range 27–46 years) and in 15 ALS subjects whose mean age was 59 (range 44–75 years). An additional 6 ALS subjects were studied for other aspects of MUE analysis. ALS subjects were followed in the Motor Neuron Disease Clinic and fulfilled clinical and electrophysiologic criteria for upper and lower motor neuron loss.\textsuperscript{2} The two estimates (test and retest) were separated by 1 day to 4 months in normal subjects (median 8 days), and by 1 to 7 days in ALS subjects (median 1 day).

The spike-triggering averaging technique was modified from that previously described.\textsuperscript{6} A 7 \times 80 mm silver strip electrode was placed transversely across the surface of the biceps-brachial muscle at the motor point. This site was chosen for consistency of electrode placement during reproducibility testing. A reference disc electrode was placed over the olecranon process. The maximum CMAP from the biceps-brachialis muscle group was obtained by electrical stimulation over the musculocutaneous nerve, avoiding activation of radial or median innervated muscles.

An intramuscular recording electrode (single-fiber macro-EMG or concentric needle type) was introduced into the biceps muscle 3–5 cm distal to the recording strip electrode, and used to isolate the triggering motor unit spike potentials which were generated by weak voluntary muscle contractions. The stability of the spike potentials when the surface motor unit action potential (S-MUP) was recorded was monitored by a raster display of the spike potentials. Each S-MUP obtained was the average of 150 to 250 responses. The corresponding triggered motor unit spike potentials were obtained at five different depths and from three electrode penetrations at medial, central, and lateral sites across the muscle. No S-MUP was discarded unless it was identical in shape and peak-to-peak amplitude to the preceding potential.

The MUE was calculated as the ratio of the peak-to-peak CMAP amplitude divided by the mean peak-to-peak S-MUP amplitude. For routine MUE calculations, 15 S-MUPs were obtained from each muscle and averaged. In other phases of the study, additional S-MUPs were obtained at each penetration site and 25 S-MUPs were averaged.

Test–retest reproducibility was assessed in two ways. The relative difference was calculated as the absolute value of the difference between test and retest values divided by their mean value and expressed as a percentage. The correlation coefficient between test and retest values was calculated to determine how closely the test value predicted the retest value. The magnitude of the correlation coefficients can be disproportionally influenced by data points with large values. The disproportionate influence can be reduced by a square root transformation,\textsuperscript{20} and this was applied where indicated. Statistical comparisons, except where noted, were made by the Wilcoxon signed-rank test because the distribution of many of the data sets appeared, on inspection, to be highly skewed.

**RESULTS**

The initial or test MUE values from normal subjects (mean 553, range 235–1580) were significantly greater ($P = 0.005$) than those from ALS subjects (mean 149, range 53–343), and there was little overlap in the distributions (Table 1, Fig. 1). The lowest MUE test value from normal subjects (235) was taken as the lower limit of normal.\textsuperscript{11,25}

When test–retest MUE values were compared (Table 1), there were no significant differences for normal subjects ($P = 0.72$) or ALS subjects ($P = 0.27$). In normal subjects, the average test–retest relative difference was 45.3%, while in ALS subjects the relative difference was 32.6%. In normal subjects, the test–retest correlation coefficient was $r = 0.86$, but a plot of test–retest MUE values revealed scatter and one influential outlying point (Fig. 2). The correlation coefficient fell to $r = 0.07$ after removal of that point (Table 1). In ALS subjects, the correlation coefficient was $r = 0.54$. A plot of test–retest values also displayed scatter, but the scatter was prominent only among subjects with larger MUE values (Fig. 2). When 4 ALS subjects with MUE values lying within the normal range (>235) were removed the correlation coefficient rose to $r = 0.65$ (Table 1). The influence of large MUE values in ALS subjects was also assessed.
Table 1. Test–retest reproducibility of MUE, CMAP, and S-MUP values in normal and ALS subjects.

<table>
<thead>
<tr>
<th></th>
<th>Mean (n)</th>
<th>SD</th>
<th>Range</th>
<th>P-value</th>
<th>Relative difference (%)</th>
<th>Correlation coefficient full (restricted) data*</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>MUE</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Normal subjects</td>
<td>(10)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test</td>
<td>553</td>
<td>380</td>
<td>235–1580</td>
<td>0.72</td>
<td>45.3</td>
<td>0.86 (0.07)</td>
</tr>
<tr>
<td>Retest</td>
<td>591</td>
<td>481</td>
<td>211–1816</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ALS subjects</td>
<td>(15)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test</td>
<td>149</td>
<td>87</td>
<td>53–343</td>
<td>0.27</td>
<td>32.6</td>
<td>0.54 (0.65)</td>
</tr>
<tr>
<td>Retest</td>
<td>154</td>
<td>69</td>
<td>74–298</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>CMAP (µV)</strong></td>
<td>(10)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Normal subjects</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test</td>
<td>11,609</td>
<td>3919</td>
<td>5053–17,380</td>
<td>0.88</td>
<td>20.3</td>
<td>0.57</td>
</tr>
<tr>
<td>Retest</td>
<td>11,938</td>
<td>3207</td>
<td>6395–16,895</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ALS subjects</td>
<td>(15)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test</td>
<td>7257</td>
<td>2971</td>
<td>2595–11,495</td>
<td>0.91</td>
<td>21.4</td>
<td>0.84</td>
</tr>
<tr>
<td>Retest</td>
<td>7228</td>
<td>2943</td>
<td>3563–13,050</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>S-MUP (µV)</strong></td>
<td>(10)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Normal subjects</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test</td>
<td>26</td>
<td>14</td>
<td>11–63</td>
<td>0.54</td>
<td>44.2</td>
<td>0.20</td>
</tr>
<tr>
<td>Retest</td>
<td>28</td>
<td>12.7</td>
<td>9–44</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ALS subjects</td>
<td>(15)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test</td>
<td>59</td>
<td>35</td>
<td>23–143</td>
<td>0.15</td>
<td>23.8</td>
<td>0.81</td>
</tr>
<tr>
<td>Retest</td>
<td>54</td>
<td>27</td>
<td>19–98</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

by performing a square root transformation. When correlation coefficients were compared, the coefficient rose from $r = 0.54$ with raw data to $r = 0.62$ with transformed data (Table 1), supporting poorer correlations among large MUE values.

Further analysis of test–retest variability was made by reviewing the CMAP and mean S-MUP amplitude values from which the MUE was calculated. The average test–retest relative difference for CMAP amplitude was similar among normal (20.3%) and ALS subjects (21.4%) (Table 1). There was no significant difference between test–retest CMAP values ($P = 0.88$, normal subjects; $P = 0.91$, ALS subjects). There was, however, greater scatter of the test–retest CMAP data points and lower correlation coefficients for normal subjects ($r = 0.57$) than for ALS subjects ($r = 0.84$).

The distributions of S-MUPs recorded from both normal and ALS subjects were skewed and more broad for ALS subjects (Fig. 3). The average test–retest S-MUP values were not significantly different among normal subjects ($P = 0.54$) and ALS subjects ($P = 0.15$) (Table 1). The distributions of S-MUP average values recorded in each subject were tested to determine if test–retest samples were homogenous. Significant test–retest differences ($P < 0.05$) in S-MUP amplitudes were found in 7 of 10 normal subjects, but in only 2 of 15 ALS subjects.

The influence of the number of S-MUP amplitude values averaged for use in the MUE calculation was investigated in two ways. In the first, 15 different MUE values were calculated and plotted as the number of S-MUPs averaged for use in the denominator increased sequentially from one to 15 (Fig. 4). No trends of increasing or decreasing MUE values were apparent for normal or ALS subjects. In several ALS subjects, shifts in MUE values were noted which were associated with repositioning of the spike-triggering electrode (dif-
FIGURE 2. Scattergrams of test–retest MUE values in normal subjects (left) and ALS subjects (right). Lower limit of normal MUE range (235) marked by arrows for ALS subjects. See text and Table 1 for correlation coefficients.

different penetrations) in the muscle, suggesting that there are regional differences in the size of motor units, but the effects on MUE were minimal. None of the MUE values from ALS subjects strayed into the normal range (>235) after 10 S-MUPs were averaged, and a comparison of MUE values calculated from averages of 10 and 15 S-MUP values showed no significant difference (P = 0.17, normal subjects; P = 0.15, ALS subjects). The sequential MUE values were also plotted as a percentage of the final MUE value (when the denominator was 15) to determine how quickly the values converge (Fig. 4). There was substantial variability when few S-MUPs were averaged, but convergence occurred after 10 values were averaged.

In the second method, carried out in 6 different ALS subjects, additional S-MUPs were recorded in each of the three electrode penetrations for a total sample of 25 S-MUPs compared to the routine sample of 15 S-MUPs. The MUEs calculated from the average of 15 S-MUPs were compared to those calculated from the expanded average of 25 S-MUPs, and there was no significant difference (Student’s t-test, P = 0.46).

DISCUSSION

The ranges of MUE values obtained in this study for normal and ALS subjects are similar to those reported for the same technique in the same muscle group,6,25 and consequently, the lower limit of normal is similar.25

Comparisons of MUE reproducibility are limited by differences in MUE techniques, muscles tested, and methods of expression of reproducibility. Within these limitations, McComas and co-workers,16 using the incremental stimulation technique, reported that, when motor unit counts were estimated twice in the same day in the extensor

FIGURE 3. Frequency distributions of S-MUP amplitude values of normal subjects (left) and ALS subjects (right).
digitorum brevis muscle of normal subjects, the larger value exceeds the smaller by an average of 11.4% and, when retested on separate days, by an average of 33.7%. This value is close to the 45.3% relative difference found in the current study (Table 1), using a different technique in a different muscle group. Stein and Yang compared five MUE techniques, including the two listed above, in thenar muscles of normal subjects and found no significant differences in MUE values.

The CMAP is recorded in a similar manner in all MUE techniques, and common concerns which may affect accuracy and reproducibility include volume conduction from nearby muscles and supramaximal excitation to activate all motor units. The major factor governing MUE accuracy and reproducibility is the determination of the S-MUP value to use in the calculation. All methods of determination involve a sampling of the motor unit population. For the spike-triggered averaging technique, a sample of S-MUPs is obtained and a true mean S-MUP value is calculated. For the incremental stimulation technique, a mean contributing S-MUP value is determined. The size distributions of normal motor units innervating various muscles are not known, but an estimate can be inferred from fiber diameter histograms calculated from human ventral roots, and it is likely that no S-MUP sampling technique is representative.

Reproducibility was higher in ALS subjects with low motor unit counts. Under these conditions, the chance of sampling the same motor unit during retesting will be greater. When motor unit counts are high, the probability of sampling the same motor unit during retesting will be lower, and the inherent variability of the population of motor units will influence reproducibility. This can account for the finding in the current study that 7 of 10 normal subjects had significant differences between test-retest S-MUP amplitude samples compared with 2 of 15 ALS subjects. In support, Brandstater et al. studying limb muscles and Delbeke, studying facial muscles, report less test-retest variability (lower standard deviations) when subjects who had reduced numbers of motor units were studied.

In ALS, two factors will broaden the distribution of S-MUP amplitude values in the spike-triggered averaging technique (Fig. 3). During denervation there is loss of low- and high-threshold motor units and, during reinnervation, the size of motor units changes. It has been suggested that, under these conditions, larger sample
sizes are necessary to accurately estimate the S-MUP value.15

Few studies have examined the effect of S-MUP sample size on the MUE calculation. In the initial description of the spike-triggered averaging technique at least 10 S-MUPs were averaged.6 In the manual version of the incremental nerve stimulation technique discernment of more than 12 steps becomes difficult16 and, in the commercially available automated version of this technique, up to 20 steps can be extracted.11 In the current study, differences in MUE values were not clinically important when more than 10 S-MUPs were averaged. This is supported by finding no statistically significant differences in MUE values when 10, 15, or 25 S-MUPs were averaged (Fig. 4).

Other techniques to optimize the S-MUP sample have been proposed. In the automated version of the incremental nerve stimulation technique, averaging three or more MUE trials reduces an error statistic.8,11 In a briefly reported technique, based on the Poisson distribution to calculate the average S-MUP contribution, S-MUP samples are obtained until their standard error is less than 10% of the mean value (Daube, personal communication). These optimization techniques are attractive but require online computer analysis and proprietary software.

In summary, test-retest reproducibility of the spike-triggered averaging MUE technique is similar to that reported in normal subjects for the incremental stimulation technique. The variability of MUE when motor unit counts are high may limit the technique as a diagnostic tool in situations of equivocal denervation. However, the reduced variability when counts are low support the use of MUE in conditions of more marked denervation. Recent reports demonstrate the sensitivity of MUE. Preliminary information on the rate of motor unit loss in ALS indicates a nonlinear pattern.7,9 A study of the rate of motor unit loss in subjects with old poliomyelitis shows that the rate is not accelerated, but is similar to the rate of loss with aging in normal subjects.14 In this capacity, MUE may be useful to assess changes in the rate of motor unit loss during treatment trials of ALS subjects.

REFERENCES