ELSEVIER

Clinical Neurophysiology 117 (2006) 596–603

www.elsevier.com/locate/clinph

Within-subject reliability of motor unit number estimates and quantitative motor unit analysis in a distal and proximal upper limb muscle

Shaun G. Boe [a], Daniel W. Stashuk [b], Timothy J. Doherty [a,c,*]

[a] *School of Kinesiology, The University of Western Ontario, Ont., Canada*
[b] *Department of Systems Design Engineering, University of Waterloo, Ont., Canada*
[c] *Departments of Clinical Neurological Sciences and Rehabilitation Medicine, The University of Western Ontario, Ont., Canada*

Accepted 30 October 2005
Available online 25 January 2006

Abstract

Objective: To establish within-subject reliability of motor unit number estimates (MUNEs) and quantitative MU analysis using decomposition-based quantitative electromyography (DQEMG).

Methods: Following the acquisition of a maximum M-wave, needle and surface-detected EMG signals were collected during contractions of the first dorsal interrosseous (FDI) and biceps brachii (BB). DQEMG was used to extract motor unit potential (MUP) trains and surface-detected MUPs associated with each train, the mean size of which was divided into the maximum M-wave to obtain a MUNE. Retests were performed following the initial test to evaluate reliability.

Results: Subjects test-retest MUNEs were highly correlated ($r=0.72$ FDI; 0.97 BB) with no significant differences between test and retest MUNE values ($P>0.10$). Ninety-five percent confidence intervals were calculated to establish the range of expected retest MUNE variability and were ± 41 MUs for the FDI and BB. Quantitative information pertaining to MU size, complexity and firing rate were similar for both tests.

Conclusion: MUNEs and quantitative MU data can be obtained reliably from the BB and FDI using DQEMG in individual subjects.

Significance: Establishing within-subject reliability of MUNEs and quantitative MU analysis allow clinicians to longitudinally follow changes in the MU pool of individuals with disorders of the central or peripheral nervous system in addition to assessing their response to treatments.

Keywords: Motor unit number estimation; Quantitative EMG; Reliability; Electromyography; Motor unit

1. Introduction

Motor unit number estimates (MUNEs) and quantitative motor unit (MU) analysis performed using decomposition-based quantitative electromyography (DQEMG) provide clinically useful information pertaining to the physiological characteristics of individual MUs and the size of the underlying MU pool within a given muscle group. Taken together, this information may enable clinicians to better characterize the extent of MU loss and subsequent reorganization of the MU pool in response to disorders of the central and peripheral nervous systems, and to follow the natural history and response to treatment of these disorders.

To provide value as a clinical tool, the results obtained with DQEMG must be reliable so that changes from test to test may be interpreted as resulting from some process other than the imprecision of the technique. Previously, DQEMG derived MUNEs of the thenar muscle group were shown to be highly reliable across a population of healthy younger adults while other studies using these methods have attempted to characterize the numbers and characteristics of MUs within varying muscles (Boe et al., 2004; Doherty and Stashuk, 2003; McNeil et al., 2005a). While these studies have contributed significantly to the advancement of the technique, they have not adequately addressed clinical

* Corresponding author. Address: London Health Sciences Centre, University Hospital, 339 Windermere Road, London, Ont., Canada N6A 5A5. Tel.: +1 519 663 3337l; fax: +1 519 663 3328.
E-mail address: tim.doherty@lhsc.on.ca (T.J. Doherty).

1388-2457/$30.00
doi:10.1016/j.clinph.2005.10.021

reliability as they have focused on the reliability of the method for a population, as opposed to the expected variability on test-retest for any one subject.

Consequently, it is essential to examine the reliability of the data obtained using DQEMG within individual subjects so that it may be used longitudinally to follow changes within a given subjects MU pool. In order to achieve this, a range of MUNE values must be determined that takes into account the degree of variability associated with the DQEMG technique. In a manner similar to those employed previously (Simmons et al., 2001), the calculation of confidence intervals for an individual subject's predicted MUNE value has been used to allow for the identification of those changes to the MU pool that fall outside of those expected to occur due to methodological variability, and therefore, result from disorders of the central or peripheral nervous system.

Thus, the purpose of the current study was to determine the within-subject reliability of MUNEs and quantitative MU analysis for the first dorsal interosseous (FDI) and biceps brachii (BB) muscles performed using DQEMG, and to establish 95% confidence intervals around the predicted MUNE values of individual subjects. The FDI and BB muscles were chosen due to their potential differences in force production strategies and the accessibility of their nerve supply, which allows for the acquisition of a maximum M-wave and subsequent calculation of a MUNE. Lastly, it was important to establish the reliability of MUNEs and quantitative MU analysis in muscles that represent different segments of the cervical cord, particularly for future studies of patients with motor neuron disease who may present with disease onset in different segments.

2. Methods

2.1. Subjects

Ten healthy subjects aged 27 ± 6 years volunteered to take part in the study. All gave informed consent and our institutional review board approved the study.

2.2. Force measurement

For the FDI muscle, subjects were seated during data collection with their right arm pronated and placed in a custom-made force dynamometer. In order to isolate the action of the FDI muscle, the thumb was stabilized with a metal brace at 90° extension and the lateral three digits separated from the second digit with a divider, and immobilized with a medium density sponge placed over the digits and secured with a Velcro strap. Additional straps placed just distal and proximal to the wrist joint line secured the forearm and hand position. The isometric abduction force exerted by the FDI was measured in Newtons (N) with a force transducer (Model FT-10; Grass-Telefactor, West-Warwick, RI) that was anchored to the device and aligned with the proximal interphalangeal joint of the second digit. The output from the force transducer was amplified (Model CP 122 AC/DC Amplifier; Grass-Telefactor, West-Warwick, RI) and converted to digital format by a 12-bit converter (CED model 1401 Plus, Cambridge Electronic Design, Cambridge, UK) at a sampling rate of 500 Hz and displayed on an analog oscilloscope (Model 5111A storage oscilloscope; Tektronix Inc., Beaverton, OR) placed in front of the subject.

The protocol used to measure BB force output is similar to a previously reported study (Klein et al., 2001). Subjects were supine on a padded table and the right arm placed in a custom-made force dynamometer. The legs were supported on a padded wooden box, with the hip and knee joints flexed to 90° and the right shoulder secured with a padded metal brace. The box and brace prevented the torso from sliding during contractions. The elbow joint was flexed 90° and placed in a padded cup with the forearm fully supinated. The wrist and fingers were prevented from flexing during contraction by a plastic splint that was strapped to the back of the wrist and hand. The ventral aspect of the wrist was secured with a strap to a padded curved bar (11×5.2 cm) that had a strain gauge attached (model SST-700-100A, ASTechnology, Haliburton, Ont., Canada). The output from the strain gauge was amplified (Neurolog, models NL 107, and NL 126, Digitimer, Welwyn Garden City, Hertfordshire, UK), and converted to digital format by a 12-bit converter (CED model 1401 Plus, Cambridge Electronic Design, Cambridge, UK) at a sampling rate of 500 Hz and displayed on an analog oscilloscope (Model 5111A storage oscilloscope; Tektronix Inc., Beaverton, OR) suspended above the subject.

The force signals for both the FDI and BB were analyzed off-line using a commercially available software package (Spike 2 v. 4.5; Cambridge Electronic Design, Cambridge, UK).

2.3. Electromyographic data collection

The DQEMG method and associated algorithms as described in detail elsewhere were used (Boe et al., 2004; Doherty and Stashuk, 2003). Electromyographic signals were acquired using DQEMG software on the Neuroscan Comperio (Compumedics Medical Systems, El Paso, TX). Intramuscular signals were recorded with a commercially available, disposable concentric needle electrode (Model N53153; Teca Corp., Hawthorne, NY) with a bandpass of 10 Hz–10 kHz, while surface signals were recorded with a bandpass of 5 Hz–5 kHz using self-adhering electrodes (Kendall-LTP, Chicopee, MA). For the FDI muscle, a full size electrode was cut in strips ($1 \text{ cm} \times 3 \text{ cm}$) and the active electrode located over the motor point of the muscle with the reference electrode located over the first metacarpophalangeal joint. For the BB muscle, full size electrodes ($2 \text{ cm} \times 3 \text{ cm}$) were used with the active electrode located

over the motor point of the muscle and the reference electrode located over the distal tendon. A full size electrode served as a ground for both the FDI (dorsal aspect of the hand) and BB (forearm just distal to the elbow crease) measurements.

2.4. Experimental protocol

The experimental protocol was similar for both the FDI and BB muscles. Prior to placement in the appropriate force measurement device, the maximum M-wave was elicited with supramaximal stimulation of either the ulnar nerve at the wrist (FDI) or the musculocutaneous nerve at the axilla (BB). Markers indicating negative onset, negative peak, negative-peak duration, and positive peak were automatically positioned. Following a visual check of the markers (and manual adjustments if required), size-related parameters of the M-wave including negative-peak amplitude were automatically calculated.

Following placement in the appropriate dynamometer, subjects were instructed to perform a 4-s maximal voluntary contraction (MVC) and were aided in doing so by visual feedback from the force record on the oscilloscope and strong verbal encouragement from the examiner. The peak of the MVC was determined and a number corresponding to 10% of this value was marked on the oscilloscope. Subsequent contractions were performed at approximately this level (see results for force data).

A concentric needle electrode was then inserted into the muscle of interest just proximal or distal to the active surface electrode. Subjects were then asked to minimally contract the muscle isometrically while the needle position was adjusted in order to minimize the rise-times of the MUPs of the first 2–3 recruited MUs. With the needle manually maintained in a stable position by the examiner, the subject was instructed to increase the contraction force to the desired percent of MVC. At this level of MVC, additional MUs were recruited beyond those active during the minimal contractions described above, thereby increasing the possibility of identifying multiple motor unit potential (MUP) trains during an individual contraction. If the needle signal was of poor quality based on visual inspection, the needle was repositioned and the process repeated to ensure adequate signal quality. Each isometric contraction lasted for 30-s. Subjects were instructed to maintain contractions of consistent force throughout the 30-s period and were aided in doing so by the presence of a target line displayed on the oscilloscope. In order to decrease the chances of sampling the same MUs from different contractions, the needle position was adjusted between contractions to collect from superficial, intermediate, and deep portions of the muscle. Contractions were repeated until a minimum of 20 motor unit potential (MUP) trains were obtained (see results for contraction data). If more than three contractions were required, as was often the case, a second needle insertion site was undertaken. Following needle-detected signal decomposition and analysis, the MUP trains and needle and surface-detected MUPs (S-MUPs) were reviewed with regard to their acceptability based on criteria that has been previously reported (Boe et al., 2005). All MUP trains and needle and surface-detected MUPs that met the inclusion criteria were included in the subsequent data analysis and reported in the results section.

For the retest portion of the study, the repeat study was always performed a minimum of 2 days and up to 7 days following the initial study. All tests and subsequent review of the MUP trains and needle and surface-detected MUPs were performed by the same examiner (SB). The initial surface electrode positions were not marked on the skin, and data analysis was completed only following collection of both the test and retest data.

2.5. Statistics

Mean values along with their standard deviations are presented throughout. Comparison between individuals test and retest values were made using standard pairwise *t* tests. Additionally, 95% confidence intervals were calculated for the predicted true scores, see Eq. (1), of each participant's MUNE using the standard error of measurement (Pedhazur, 1991). Lastly, the mean difference percentage was calculated as the absolute difference between the test and retest MUNE values divided by their mean value in order to allow for comparison to previous studies of MUNE reliability. The a priori significance level was established at $P<0.10$.

$$T = (1 - r_{xx})M + r_{xx}X \quad (1)$$

where T, predicted true score; r_{xx}, correlation coefficient of the observed test-retest values; M, mean of the observed test values; and X, subjects observed test value.

3. Results

3.1. Biceps brachii force, MUP size and firing rate

BB MVC values were similar, with mean values of 355.9 ± 41.3N and 337.2 ± 41.9N for the test and retest sessions respectively. For the 30-s isometric contractions, subjects produced an average of $10.2 \pm 2.1\%$ of their MVC value for the initial test and $11.1 \pm 2.5\%$ of their MVC value during the retest.

During the test and retest sessions, we sampled 34 ± 5 and 34 ± 8 MUs per subject from 7 ± 2 and 7 ± 1 contractions respectively with mean MU identification rates of $64.4 \pm 7.1\%$ and $63.5 \pm 6.7\%$. Despite comparable mean firing rates between the test and retest (12.1 ± 1.1 Hz and 11.4 ± 1.0 Hz), a significant difference was observed between the firing rate values of individual subjects.

Table 1
Individual subjects mean needle and surface-detected MUP and maximum M-wave size and MUNE values for the biceps brachii

Parameter	Needle-detected MUP		Surface-detected MUP		M-wave		Mune		
	Peak–peak voltage (μV)		Negative-peak amplitude (μV)		Negative-peak amplitude (mV)				
Subject	Test	Retest	Test	Retest	Test	Retest	Test	Retest	Predicted ± CI
1	353.3	294.6	50.3	57.9	12.1	12.1	256	223	256 ± 41
2	394.2	426.1	73.5	84.4	12.4	12.6	182	185	185 ± 41
3	562.3	468.4	37.4	46.4	11.1	11.7	357	279	355 ± 41
4	452.9	297.2	78.7	82.7	15.5	17.4	211	214	213 ± 41
5	398.3	328.2	91.3	93.1	15.0	15.4	170	166	173 ± 41
6	315.7	255.0	32.4	27.4	13.1	13.1	547	569	539 ± 41
7	230.7	261.3	50.4	67.0	7.3	8.8	159	140	162 ± 41
8	331.1	342.2	59.1	62.8	11.7	11.7	216	208	218 ± 41
9	318.6	279.1	33.8	36.7	11.9	10.2	395	352	392 ± 41
10	363.4	455.2	48.7	46.7	9.1	9.0	223	230	224 ± 41
Mean	372.0	340.7	55.5	60.5	11.9	12.2	272	257	272 ± 41
Std dev	89.4	80.5	20.0	21.7	2.4	2.7	124	125	121

Individual subjects predicted MUNE values ± a 95% confidence interval of 41 MUs is in the column to the far right.

Needle-detected MUP size based on peak–peak voltage was similar for both the test and retest with mean values of 372.0 ± 89.4 μV and 340.7 ± 80.5 μV. Statistical analysis of the size distributions of needle-detected MUPs of individual subjects revealed no significant differences between the test and retest (Table 1). Additionally, the number of turns and phases of the needle-detected MUPs were the same for the test and retest, with a mean value of 3 turns and 3 phases for each session. Furthermore, there were no significant differences between the numbers of turns and phases between the test and retest of individual subjects. Needle-detected MUP duration was also similar between tests, with mean values of 9.8 ± 2.5 ms and 9.3 ± 1.4 ms with no significant differences observed between individuals test-retest values. S-MUP size based on negative-peak amplitude was also similar for the test and retest with mean values of 55.5 ± 20 μV and 60.5 ± 21.7 μV. Despite the similarities in mean values, a statistically significant difference was observed between the size distributions of the S-MUPs of individuals' test-retest values (Table 1).

3.2. Biceps brachii maximum M-wave and MUNE

Maximal M-wave values (negative-peak amplitude) were similar for both tests with mean values of 11.9 ± 2.4 mV and 12.2 ± 2.7 mV. Additionally, no significant differences were observed between individual test and retest maximum M-wave values (Table 1). BB MUNE values (negative-peak amplitude) were highly reliable, displaying a strong positive relationship ($r = .97$) and similar mean MUNE values for the test and retest with values of 272 ± 174 and 257 ± 125. Furthermore, no significant differences were observed for individual BB MUNE values between the test and retest (Table 1), with a mean difference (%) between subjects test-retest MUNE values of 7.9%. Lastly, in order to identify the variability expected within the methodology itself, a 95% confidence interval was calculated for individuals' predicted BB MUNE values and determined to be ±41 MUs (Table 1, Fig. 1).

3.3. First dorsal interosseous force, MUP size and firing rate

FDI MVC values were similar for both the test and retest with mean values of 26.5 ± 5.2N and 24.3 ± 4.5N. During the 30-s isometric contractions, subjects produced an average of 9.6 ± 3.0% and 8.1 ± 2.6% of their MVC value for the test and retest sessions, respectively.

Throughout the test and retest sessions we sampled 33 ± 5 and 32 ± 4 MUs per subject from 7 ± 2 and 7 ± 1 contractions with mean MU identification rates of 56.3 ± 5.5% and 59.2 ± 5.3%. No significant differences were observed between the test and retest firing rates of individual subjects with similar mean firing rates being recorded (13.0 ± 1.6 Hz and 12.5 ± 1.3, test-retest, respectively). Test-retest needle-detected MUP size based on peak-peak voltage were also similar in the FDI with mean test-retest values of 469.1 ± 118.4 μV and 446.6 ± 96.0 μV.

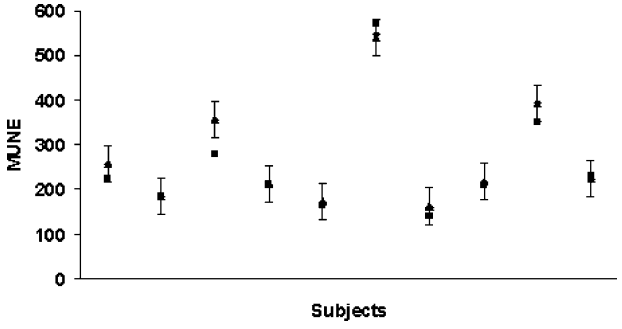

Fig. 1. Individual subjects test (diamond), retest (square) and predicted (triangle) MUNE values for the biceps brachii muscle. Bars represent a 95% confidence interval of 41 MUs centered on the predicted MUNE value. Note all subjects retest MUNE values are located within the 95% confidence interval with the exception of subject 3.

Table 2
Individual subjects mean needle and surface-detected MUP and maximum M-wave size and MUNE values for the first dorsal interosseous

	Needle-detected MUP		Surface-detected MUP		M-wave		Mune		
Parameter	Peak–peak voltage (μV)		Negative-peak amplitude (μV)		Negative-peak amplitude (mV)				
Subject	Test	Retest	Test	Retest	Test	Retest	Test	Retest	Predicted ± CI
1	582.5	528.1	147.5	102.9	17.2	17.8	130	187	131 ± 41
2	443.6	574.2	229.0	181.2	15.5	16.0	78	90	90 ± 41
3	460.8	424.4	88.3	79.6	14.7	14.0	204	187	189 ± 41
4	487.9	388.5	113.2	156.4	15.8	16.9	137	111	136 ± 41
5	726.6	561.0	105.2	119.2	10.5	12.7	112	134	117 ± 41
6	489.9	472.0	112.5	83.7	10.9	11.2	116	150	120 ± 41
7	370.2	311.0	97.0	100.4	11.3	12.8	131	146	132 ± 41
8	325.6	298.8	167.4	152.9	18.8	18.6	140	152	139 ± 41
9	347.1	486.1	66.0	71.0	10.9	11.2	213	179	196 ± 41
10	456.5	422.1	155.1	145.8	11.2	10.9	83	84	94 ± 41
Mean	469.1	446.6	128.1	119.3	13.7	14.2	134	142	134 ± 41
Std dev	118.4	96.0	47.4	37.8	3.1	2.9	44	38	35

Individual subjects predicted MUNE values ± a 95% confidence interval of 41 MUs is in the column to the far right.

Likewise, analysis of the size distributions of needle-detected MUP size of individual subjects between the test and the retest revealed no significant differences (Table 2). Additionally, the same number of turns and phases of the needle-detected MUPs were observed for the test and retest, with mean values of 4 turns and 3 phases, respectively. Furthermore, no significant differences were observed between the number of turns and phases for the individual subjects test and retest. Needle-detected MUP duration was also similar between tests with mean values of 9.4 ± 1.3 ms and 9.2 ± 1.8 ms with no significant differences observed between individuals test-retest values. S-MUP size based on negative-peak amplitude was comparable between the test and the retest with mean values of 128.1 ± 47.4 μV and 119.3 ± 37.8 μV. Following statistical analysis, it was concluded that there were no significant differences between test-retest size distributions of the S-MUPs of individual subjects (Table 2).

3.4. First dorsal interosseous maximum M-wave and MUNE

Maximal M-wave values (negative-peak amplitude) were similar for both the test and retest with mean values of 13.7 ± 3.1 mV and 14.2 ± 2.9 mV. Although the mean values are comparable, there was enough variability present in the individual maximum M-wave values to reach significance (Table 2). Despite this finding, FDI MUNE values (negative-peak amplitude) were reliable, displaying a strong positive relationship ($r = .72$) and similar mean MUNE values of 134 ± 44 and 142 ± 38 for the test–retest, respectively. Additionally, statistical analysis of individual subjects MUNE values revealed no significant differences between the test and retest (Table 2), with a mean difference (%) of 16.1% between the two tests. Similar to the BB, a 95% confidence interval was calculated for individuals' predicted FDI MUNE values and determined to be ±41 MUs (Table 1, Fig. 2).

4. Discussion

This study has demonstrated that quantitative information pertaining to the physiological characteristics and size of the MU pool, including MUNEs, can be obtained reliably in individual subjects using DQEMG in the FDI and BB, enhancing the ability to follow the natural history and the response to treatments of disorders of the peripheral or central nervous system. The importance of establishing within-subject reliability is evident in the sizes of the standard deviations of the mean values of the current study including MUNE and needle and surface-detected MUP size (Tables 1 and 2). These large standard deviations suggest that a great deal of physiological variability is present between subjects. Therefore, in any longitudinal study, it is inappropriate to follow group mean scores and thus necessary to monitor changes in individual subjects.

As previously outlined, one of the purposes of this study was to determine a range of values around an individual subjects MUNE value that would take into account the

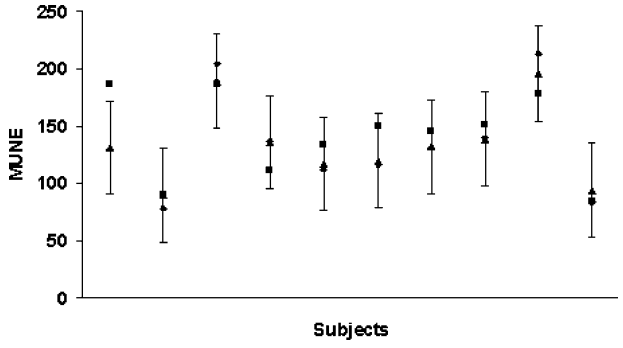

Fig. 2. Individual subjects test (diamond), retest (square) and predicted (triangle) MUNE values for the first dorsal interosseous muscle. Bars represent a 95% confidence interval of 41 MUs centered on the predicted MUNE value. Note all subjects retest MUNE values are located within the 95% confidence interval with the exception of subject 1.

Table 3
Summary of MUNE reliability studies

Investigator	Method	Muscle	Subjects	Correlation coefficient (r)	Mean difference (%)
Bromberg	Spike-triggered averaging	Biceps-brachialis	10 Control subjects, 36 years (27–46 years)	0.86	45.3
Felice	Multiple point stimulation (MPS)	Thenar	16 Control subjects, 39 years (24–52 years)	0.85	17.0
Shefner et al.	Statistical modified statistical	Adductor digiti minimi	18 Control subjects, 24–67 years	0.44 (statistical) 0.57 (modified)	23.0 statistical 14.0 modified
Lomen-Hoerth and Olney	MPS and statistical	Hypothenar	20 Control subjects, 24–53 years	0.90 (MPS) 0.98 (statistical)	–
Doherty and Brown	MPS	Thenar	17 Control subjects, 21–81 years	0.88	–
Current study	DQEMG	BB and FDI	10 Control subjects, 27 years (20–40 years)	0.97 (BB) 0.72 (FDI)	7.9 (BB), 16.1 (FDI)

variability of the method. The application of confidence intervals addresses this purpose, allowing clinicians to conclude with 95% confidence that when a given MUNE is calculated its true score will fall within this range of MUNE values. Therefore, in any longitudinal or intervention study, one can conclude that values falling outside of this range are due to disease progression or the intervention as opposed to methodological error. These findings are highlighted in Figs. 1 and 2, which display subjects individual test and retest MUNE values in addition to their predicted (true score) MUNE values $\pm$ the 95% confidence interval of 41 MUs. With the exception of two subjects (subject 3, Fig. 1 and subject 1, Fig. 2), all subjects retest values were within the range of the predicted (true score) MUNE $\pm$ the 41 MUs.

The reliability and 95% confidence interval range for the BB and FDI muscles displayed in the current study are similar to those reported previously. First, reliability, as determined by mean difference (%) and correlation coefficient (r) between test and retest MUNE values in the current study, compare favorably to previous studies that have applied various MUNE methods in several muscle groups (Table 3) (Bromberg, 1993; Doherty and Brown, 1993; Felice, 1995; Lomen-Hoerth and Olney, 2000; Shefner et al., 1999) Secondly, the 95% confidence interval range of $\pm$41 MUs for the BB and FDI muscles in the current study are similar to those reported in a previous study using different variations of the statistical MUNE technique applied to the thenar muscle group (Simmons et al., 2001). Lastly, MUNE values obtained with a similar protocol (and similar force levels) from the FDI yielded mean MUNE values that are within the range of those currently reported (Boe et al., 2004; 2005).

As this is the first study to derive MUNEs for the BB using DQEMG, it is inappropriate to directly compare the values obtained in the current study to those obtained using traditional spike-triggered averaging (STA). Perhaps as expected, the MUNE values from the current study are considerably lower when compared to the previous studies for this muscle group (Bromberg, 1993; Brown et al., 1988; Doherty et al., 1993). These lower values can be attributed to the fact that the sample of MUs from traditional STA are drawn from low-level voluntary contractions and thus is potentially biased to smaller S-MUPs. This will result in noticeably higher MUNEs than those obtained with DQEMG, which is able to sample S-MUPs from significantly higher levels of voluntary contraction and thus obtain a more diverse sample of MUPs from a wider physiological range of MUs (Boe et al., 2005).

In addition to the reliability of DQEMG derived MUNEs, the current results illustrate that quantitative information pertaining to the electrophysiological characteristics and discharge behavior of individual MUs can be obtained reliably using DQEMG. This information, including MUP amplitude, duration, complexity, firing rate and variability can be invaluable in determining the nature of a disorder (for example neuropathic versus myopathic MUPs), as well as understanding the extent of the reorganization of the MU pool in response to these disorders. This quantitative information, in addition to MUNE, may provide additional useful information regarding the nature of the underlying problem and its time course (for example, the presence of an acute or chronic process).

Significant differences were observed between individual subjects test–retest values for S-MUP size and MU firing rate for the BB. The differences reported for BB S-MUP size may be due to variation in electrode position as the initial electrode positions in this study were not marked, and the retest locations were determined by maximizing the amplitude of the maximum M-wave. As a result, small deviations in electrode position could potentially have contributed to the slight changes in S-MUP amplitude observed in the current study (mean difference between the test and retest of 7.8%).

Despite these small changes, the MUNE values in the current study (and in particular for the BB) were not significantly different between tests. Because MUNE is the result of dividing the mean S-MUP size into the maximum M-wave size, any small differences in electrode position between the test and retest will be reflected in both the maximum M-wave and mean S-MUP size. For this reason,

MUNE may be more appropriate to use for test-retest purposes as S-MUP size may vary due to changes in electrode position. These effects can be limited by ensuring optimal active electrode placement during serial tests by maximizing the amplitude and minimizing the rise time of the maximum M-wave.

Small differences were also observed between the mean MU firing rates of individual subjects from test to retest. Although it is clear from the current results that a similar distribution of needle-detected MUPs was sampled, it is assumed that the small mean difference (5.9%) in MU firing rates resulted from sampling a slightly different population of MUs during the retest. This assumption is based on the fact that other variables, most importantly the level of contractile force, were controlled for during the experimental protocol. Also, it should be noted that a liberal alpha level of 0.10 was employed in order to allow for a more confident evaluation of reliability. As a result, some variables reached significance despite only a small difference between the test and retest values that would not have reached significant had a more conservative alpha level (0.05) been utilized.

It has been shown in the current study that the results obtained with the DQEMG technique are reliable within individual subjects. Subsequent to this finding is the importance of discussing the sources of variability in the DQEMG technique, which include the acquisition of the maximal M-wave, variations in the position of the surface electrodes and lastly, the level of voluntary force at which the studies are performed.

Variation in the size of the maximal M-wave can alter the MUNE value, as MUNE is dependent on both the size of the maximal M-wave and the mean S-MUP. The current results show that small variations in the maximal M-wave values between subjects tests and retests did not significantly affect the MUNE values that were obtained, although they are a potential source of variability inherent in the methodology. Secondly, the current results show that electrode position can influence the size of the S-MUPs, the effect of which has been discussed above. The differences observed between mean S-MUP sizes between tests coupled with the finding of similar MUNE and maximum M-wave values suggest that S-MUP sampling is a source of variability in MUNE. This variability may be inherent in the process of S-MUP sampling rather than the actual number of S-MUPs sampled, as a greater number of S-MUPs were sampled in the current study than the number previously suggested necessary for performing MUNE (Doherty and Brown, 1993).

Previous studies examining the DQEMG technique have shown that the level of voluntary force has a significant impact on the sizes of both the needle and surface-detected MUPs obtained with DQEMG (Boe et al., 2005; McNeil et al., 2005b). At higher force levels, larger needle and surface-detected MUPs are acquired, resulting in lower MUNEs (Boe et al., 2005; McNeil et al., 2005b). Conversely, as with standard STA, sampling S-MUPs at lower levels of force will result in smaller S-MUPs and relatively higher MUNE values. As a result of this force level effect, it is necessary to monitor the level of voluntary force in order to sample a similar population of the underlying MUs within a given muscle or muscle group. This is not only important for MUNE, but also for longitudinally monitoring changes in the MU pool via the complexity of needle-detected MUPs or their firing rates. Monitoring the level of voluntary force may be realized through the use of dynamometers (as in the current study), or more efficiently, by monitoring the surface-electromyogram (e.g. its root mean square value) which has been shown to be correlated with force in the FDI (Boe et al., 2004) and BB (unpublished observation).

Based on the present results, it can be concluded that quantitative MU analysis and MUNEs performed with DQEMG can be obtained reliably from the FDI and BB muscles. The ability of DQEMG to reliably assess the response of the underlying MU pool of individual subjects to disorders of the central and peripheral nervous system in a longitudinal fashion is important for following the natural history and the response to treatments of these disorders. Additionally, by coupling the reliability of MUNE with 95% confidence intervals, clinicians can distinguish between changes in the numbers of MUs that are due to these underlying disorders as opposed to those associated with methodological variability.

References

Boe SG, Stashuk DW, Doherty TJ. Motor unit number estimation by decomposition-enhanced spike-triggered averaging: control data, test-retest reliability, and contractile level effects. Muscle Nerve 2004;29:693–9.

Boe SG, Stashuk DW, Brown WF, Doherty TJ. Decomposition-based quantitative electromyography: effect of force on motor unit potentials and motor unit number estimates. Muscle Nerve 2005;31:365–73.

Bromberg MB. Motor unit estimation: reproducibility of the spike-triggered averaging technique in normal and ALS subjects. Muscle Nerve 1993;16:466–71.

Brown WF, Strong MJ, Snow R. Methods for estimating numbers of motor units in biceps-brachialis muscles and losses of motor units with aging. Muscle Nerve 1988;11:423–32.

Doherty TJ, Brown WF. The estimated numbers and relative sizes of thenar motor units as selected by multiple point stimulation in young and older adults. Muscle Nerve 1993;16:355–66.

Doherty TJ, Stashuk DW. Decomposition-based quantitative electromyography: methods and initial normative data in five muscles. Muscle Nerve 2003;28:204–11.

Doherty TJ, Vandervoort AA, Taylor AW, Brown WF. Effects of motor unit losses on strength in older men and women. J Appl Physiol 1993;74:868–74.

Felice KJ. Thenar motor unit number estimates using the multiple point stimulation technique: reproducibility studies in ALS patients and normal subjects. Muscle Nerve 1995;18:1412–6.

Klein CS, Rice CL, Marsh GD. Normalized force, activation, and coactivation in the arm muscles of young and old men. J Appl Physiol 2001;91:1341–9.

Lomen-Hoerth C, Olney RK. Comparison of multiple point and statistical motor unit number estimation. Muscle Nerve 2000;23:1525–33.

McNeil CJ, Doherty TJ, Stashuk DW, Rice CL. Motor unit number estimates in the tibialis anterior muscle of young, old, and very old men. Muscle Nerve 2005a;31:461–7.

McNeil CJ, Doherty TJ, Stashuk DW, Rice CL. The effect of contraction intensity on motor unit number estimates of the tibialis anterior. Clin Neurophysiol 2005b;116:1342–7.

Pedhazur E, Pedhazur Schmelkin L. Measurement design and analysis: an integrated approach. Reliability. Hillsdale, NJ: Lawrence Erlbaum and Associates, 1991.

Shefner JM, Jillapalli D, Bradshaw DY. Reducing intersubject variability in motor unit number estimation. Muscle Nerve 1999;22:1457–60.

Simmons Z, Epstein DK, Borg B, Mauger DT, Kothari MJ, Shefner JM. Reproducibility of motor unit number estimation in individual subjects. Muscle Nerve 2001;24:467–73.