ABSTRACT: Clinicians who use electromyographic (EMG) signals to help determine the presence or absence of abnormality in a muscle often, with varying degrees of success, evaluate sets of motor unit potentials (MUPs) qualitatively and/or quantitatively to characterize the muscle in a clinically meaningful way. The resulting muscle characterization can be improved using automated analysis. As such, the intent of this study was to evaluate the performance of automated, conventional Means/Outlier and Probabilistic methods in converting MUP statistics into a concise, and clinically relevant, muscle characterization. Probabilistic methods combine the set of MUP characterizations, derived using Pattern Discovery (PD), of all MUPs detected from a muscle into a characterization measure that indicates normality or abnormality. Using MUP data from healthy control subjects and patients with known neuropathic disorders, a Probabilistic method that used Bayes' rule to combine MUP characterizations into a Bayesian muscle characterization (BMC) achieved a categorization accuracy of 79.7% compared to 76.4% using the Mean method ($P > 0.1$) for biceps muscles and 94.6% accuracy for the BMC method compared to 85.8% using the Mean method ($P < 0.01$) for first dorsal interosseous muscles. The BMC method can facilitate the determination of "possible," "probable," or "definite" levels for a given muscle categorization (e.g., neuropathic) whereas the conventional Means and Outlier methods support only a dichotomous "normal" or "abnormal" decision. This work demonstrates that the BMC method can provide information that may be more useful in supporting clinical decisions than that provided by the conventional Means or Outlier methods.
Muscle Nerve **41:** 18–31, 2010

PROBABILISTIC MUSCLE CHARACTERIZATION USING QEMG: APPLICATION TO NEUROPATHIC MUSCLE

L.J. PINO, PhD,[1] D.W. STASHUK, PhD,[1] S.G. BOE, MPT, PhD,[2] and T.J. DOHERTY, MD, PhD[3]

[1] Systems Design Engineering, University of Waterloo, 200 University Avenue West, Waterloo, Ontario, Canada N2L 3G1
[2] School of Kinesiology, University of Western Ontario, London, Ontario, Canada
[3] Department of Clinical Neurological Sciences, Schulich School of Medicine and Dentistry, University of Western Ontario, London, Ontario, Canada

Accepted 21 May 2009

Clinicians routinely employ the analysis of needle electromyographic (EMG) signals to aid in categorizing a muscle as either normal or abnormal when it is affected by a neuromuscular disorder. These muscle categorizations are based on a characterization of the muscle. Currently, most muscle characterizations are based on qualitative visual and auditory analysis of EMG signals and thus are more prone to subjective bias and misinterpretation than those based on less frequently used quantitative analysis. For example, a recent study authored by Kendall and Werner[1] showed that faculty and residents (blind to the underlying diagnosis of radiculopathy) using video EMG-based examinations had an overall agreement of only 46.9% with the correct categorization. Additionally, qualitative methods are less able than quantitative

Abbreviations: AAR, area-to-amplitude ratio; AMC, arithmetic mean muscle characterization; Amp, amplitude; BMC, Bayes'-rule muscle characterization; DQEMG, decomposition-based QEMG; Dur, duration; EAS, external anal sphincter; EMG, electromyographic; FDI, first dorsal interosseous; MUP, motor unit potential; PD, pattern discovery; QEMG, quantitative electromyography; SD, standard deviation; SI, Size Index; SSD, sensitivity-specificity deviation; Thick, thickness (i.e., AAR); WOE, weight of evidence

Key words: clinical decision support; electromyographic signal analysis; motor unit potential; neuromuscular disorder; pattern discovery; quantitative electromyography; QEMG

Correspondence to: L.J.Pino; e-mail: ljpino@uwaterloo.ca

Published online 18 September 2009 in Wiley InterScience (www.interscience.wiley.com). DOI 10.1002/mus.21456

methods to provide effective longitudinal comparisons within individual patients due to their subjective nature. However, the exhaustive set of statistics generated by quantitative analysis of EMG signals is difficult to distill into a useful muscle characterization that can be used to categorize a muscle. This significantly reduces the widespread routine clinical application of quantitative electromyography (QEMG). Augmenting existing QEMG techniques with a decision support system that transforms QEMG generated statistics into a concise muscle characterization that is sensitive to small changes resulting from an underlying neuromuscular disorder would assist with muscle categorization and thus promote wider clinical use of QEMG techniques. This would subsequently allow clinicians to more precisely assess the level of disease involvement and treatment effectiveness.[2]

In this study, two methods for automated interpretation of the quantitative MUP statistics of a single muscle for clinical decision support were compared: Means/Outlier analysis and Probabilistic muscle characterization. The Means/Outlier method compares the averages of sets of MUP feature values to determine if they are below or above normative limits and counts the number of outlier values.[3,4] Probabilistic muscle characterization was first introduced by Pfeiffer[5] as a method that uses Bayes' rule to characterize muscles by combining MUP characterizations. An MUP characterization is a set of conditional probabilities (one for each muscle category) calculated for an MUP detected in a muscle. An alternative Probabilistic method for combining characterizations of MUPs detected in a muscle into a characterization measure for a muscle category, introduced in this work, is to use the average across the set of MUP characterizations for that muscle category.

An unresolved problem with the conventional Means and Outlier methods is that increasing the number of features required to categorize a muscle as abnormal or increasing the thresholds to establish the limits of normative data increases specificity at the expense of reduced sensitivity.[6,7] When these methods are used it is difficult to set the number of features or thresholds to achieve a good balance between sensitivity and specificity. As such, the hypothesis of this work is that the use of a Probabilistic muscle characterization method will be more accurate and better able than Means/Outlier techniques to balance sensitivity and specificity without the need to choose parameter values such as the number of features or the means and outlier thresholds of normality.

In this work the accuracy of the Means/Outlier and two Probabilistic methods (Bayes' rule and arithmetic mean) for muscle characterization was evaluated. The MUP characterizations (i.e., conditional probabilities) used for the two Probabilistic methods were estimated using Pattern Discovery (PD).[6,8–10]

MATERIALS AND METHODS

Clinical MUP Data. In total, 904 MUPs were sampled from 29 biceps brachii muscles and 739 MUPs from 25 first dorsal interosseous (FDI) muscles of 16 healthy control subjects (aged 27 $\pm$ 4 years), and 282 MUPs from 9 biceps brachii muscles and 409 MUPs from 13 FDI muscles of 14 patients, including 9 patients (aged 52 $\pm$ 12 years) with clinically probable or definite amyotrophic lateral sclerosis (ALS) as defined by the revised El Escorial criteria[11] and 5 patients (aged 37 $\pm$ 11 years) with Charcot–Marie–Tooth Disease Type 1X confirmed via genetic testing.

A disposable concentric needle electrode (Model N53153; Teca, Hawthorne, New York) was used to acquire EMG signals during 30 s voluntary isometric contractions performed between 10% and 20% of each individual subject's maximal voluntary contraction using decomposition-based QEMG (DQEMG) on a Neuroscan Comperio (Neuroscan Medical Systems, El Paso, Texas) using a bandpass of 10 Hz to 10 kHz at a sampling rate of 31.2 kHz.[12–15] These EMG signals were decomposed into MUP templates using DQEMG.[16] DQEMG typically finds 51 isolated MUPs produced by a single motor unit, aligns them, and uses a median trimmed average to form the MUP template. Only those MUP templates whose maximum slope exceeded 0.5 V/s were accepted as clinically viable and used for training and testing in this study.[16,17]

Probabilistic Muscle Characterization. A Probabilistic muscle characterization consists of a set of muscle characterization measures, one for each category (e.g., myopathic, normal, or neuropathic), and combines the set of individual MUP characterizations estimated from the MUPs detected from the muscle under examination. An MUP characterization is a set of conditional probabilities, one for each category. In this work, only two broad categories were considered: normal and neuropathic. Therefore, an MUP characterization had two conditional probabilities: one was the conditional probability of the category being normal given the QEMG feature values of the detected MUP [$P(\text{normal} \mid \text{MUP}_i)$] and the other

the conditional probability of the category being neuropathic given the QEMG feature values of the detected MUP [P(neuropathic | MUP_i)]. The conditional probability of a category given MUP_i is estimated empirically from training data and can be thought of as the proportion of MUPs detected from muscles of the category with the same feature values as MUP_i among the set of MUPs detected from muscles of all categories with the same feature values as MUP_i. In this study PD was used to estimate MUP characterizations, because previous analysis showed it is accurate and transparent (able to explain its conclusions).[10] Details of the PD method are provided in Appendix A and in previous work.[6,8–10]

Individual MUP characterizations lack sufficient information to accurately categorize a muscle; as such, multiple MUP characterizations must be combined to provide a representative muscle characterization that can be used for muscle categorization. Given a set of MUPs detected from a muscle, the information provided by each MUP can be combined into a Probabilistic muscle characterization. An example of a Probabilistic muscle characterization is a 79% probability the muscle is neuropathic and a 21% probability the muscle is normal given the set of MUP characterizations of the MUPs detected from that muscle. In this work we examined two methods of combining MUP characterizations to calculate Probabilistic muscle characterization measures: averaging and Bayes' rule. The average muscle characterization (AMC) measure for a muscle category is calculated by simply taking the average of the set of MUP characterization values of that category as described in eq. (1):

$$AMC(category) = \frac{\sum_{i=1}^{N} PMUP_i(category)}{N} \qquad (1)$$

In previous work Pfeiffer[5] used Bayes' rule to combine a set of MUP characterizations into a muscle characterization. A Bayesian muscle characterization (BMC) measure for a muscle category is calculated as described in eq. (2) using Bayes' rule for multiple pieces of evidence, where each piece of evidence is an MUP characterization of that category, and each MUP characterization is assumed to be conditionally independent of each other[5,6,18]:

$$BMC(category) = \frac{\prod_{i}^{N} PMUP_i(category)}{\sum_{j=1}^{2} \left(\prod_{i}^{N} PMUP_i(category_j) \right)} \qquad (2)$$

where, for eqs. (1) and (2): category = muscle category (normal or neuropathic); $PMUP_i(category)$ = the MUP characterization value for that *category* of the i^{th} MUP detected from the muscle under test (i.e., P(*category* | MUP_i)); N = number of MUPs detected from the muscle under test.

Equation (2) assumes that the prior probabilities for each category are equal to obtain an unbiased muscle characterization that is based solely on the electrophysiological evidence provided by the detected MUPs. The numerator of eq. (2) is the product of all MUP characterizations for a category. The denominator is the sum of products for all categories, and it normalizes the characterization values for each category such that they sum to 1 across all categories.

Means and Outlier Analysis. Stewart et al.[4] developed and evaluated a technique where, given a set of MUP features, a muscle category is determined by comparing the mean value of each MUP feature, calculated across MUPs detected from a muscle under examination, to its corresponding normative threshold values. The Means method calculates the mean normative threshold value using all the MUPs sampled from control muscles from which greater than 15 MUPs were detected. Threshold values for each feature are calculated using sets of 15 or more MUPs detected from the corresponding muscle of a pool of control subjects. For each set of 15 or more MUPs the mean of the feature values is calculated to produce a mean value per feature per muscle. The overall mean and the standard deviation (SD) of the means are then calculated for each feature across the set of mean values per muscle. (Note: the overall mean and the SD of the means for each feature are obtained from the set of mean values of the control muscles and not across all MUPs.) The threshold values used for the Means method for each feature are defined as the overall mean ± 2 SDs of the means. For this study, muscles were categorized as neuropathic when one or more mean feature values fell above the threshold values of the normative range.

Stålberg et al.[3] developed a method for categorizing a muscle by counting the number of outliers of each MUP feature from the set of the first 20 MUPs detected from a muscle under study. An MUP feature value is considered to be an outlier if it is below or above the low and high outlier thresholds for that feature, respectively. The first 20 MUPs detected from each muscle in the pool

of control subjects are used to determine outlier thresholds for each feature. Muscles with less than 20 detected MUPs are not used for establishing outlier thresholds. For each muscle and for each feature, the set of feature values of the first 20 MUPs is sorted in ascending order, and the third lowest and the third highest value are added to a set of low and high outliers for that feature, respectively. The lower outlier threshold for a feature is the 5th percentile of the set of low outliers, and the high outlier threshold is the 95th percentile of the set of high outliers. A muscle for which there are three or more outlying MUP feature values of the same feature each above the higher outlier threshold was categorized as neuropathic.

As an additional method for muscle categorization, the Means and Outlier methods were combined as previously[7] and called the Combined method. If either the Means or Outlier method or both declares a neuropathic abnormality, then the Combined method categorizes the muscle as neuropathic.

The Means, Outlier, and Combined methods are referred to as the conventional methods.

Muscle Categorization Performance. Biceps brachii and FDI muscle data were used separately to determine performance of the methods. Performance of the Probabilistic and the conventional methods vary depending on the data used for training. Performance for all of the methods was determined across 10 different trials using randomly chosen training data. Only control muscles were selected for training the conventional methods, whereas muscles from both control and patient groups were used for training of the Probabilistic methods as per the definition of each method. Within a trial, all muscles were tested one at a time. All of the patient muscle data, except that of the muscle being categorized, were used for training of the Probabilistic methods. The technique of withholding a single item for testing while using all other data for training is a common way of testing pattern recognition techniques and is known as "jackknifing."[18] A set of control muscles were selected at random until the number of control muscle MUPs equaled the number of MUPs in the patient training data. For each trial the set of control muscle MUPs used for training for the Means and Probabilistic methods was the same. All muscles in the control data had 20 or greater MUPs except one biceps brachii muscle that had 19 MUPs. The set of control muscle MUPs used for training the Outlier method was identical to that used for training the Means and Probabilistic methods with the exception of trials that included the biceps brachii muscle with 19 MUPs. It was excluded from the Outlier method training since this method requires 20 or greater MUPs. A category label (i.e., normal or neuropathic) for each muscle under test was provided by the Probabilistic and the conventional methods. The muscle under test was categorized (or classified) by the Probabilistic methods as being of the category with the maximum muscle characterization value. The process of jackknifing was repeated for each control muscle as well. Sensitivity and specificity were calculated for each trial. All 10 trials were done this way for each feature set using different sets of training data from trial to trial to examine how different sets of training data affected the robustness of a method.

The MUP data from each muscle was used for testing of the conventional and Probabilistic methods regardless of the number of MUPs detected per muscle. Sensitivity was defined as the total number of muscles categorized as neuropathic divided by the total number of "true" neuropathic muscles (i.e., patient muscles). Specificity was defined as the total number of muscles categorized as normal divided by the total number of "true" normal, control muscles. Accuracy was defined as the average of sensitivity and specificity. Therefore, the traditional definition of accuracy *(true negatives + true positives)/(all muscles tested)* was not used, because it is biased toward the category that has the largest number of test muscles to be categorized. The traditional accuracy measure would be skewed by an unequal proportion of subjects to patients. For instance, the traditional accuracy measure would underweigh any results for the patients if there were fewer patients than controls.

The term sensitivity-specificity deviation (SSD) was defined as:

$$SSD = \sqrt{\left((A - Sens)^2 + (A - Spec)^2\right)/2} \qquad (3)$$

where A is accuracy—the mean of specificity and sensitivity; *Sens* is sensitivity; and *Spec* is specificity. SSD was used to determine how well a categorization method maximized both specificity and sensitivity.

Values for the following QEMG-MUP features were input to each of the different methods: amplitude, duration, area, number of phases, and number of turns.[16] In addition, area-to-amplitude

Table 1. Biceps brachii: average performance across all feature sets.

		# Features used for testing							
		Two		Three		Four		Five	
Method		Mean	SD	Mean	SD	Mean	SD	Mean	SD
AMC	Sens.	73.1%	9.9%	73.7%	4.9%	73.0%	3.4%	71.6%	2.8%
	Spec.	70.3%	6.9%	76.4%	4.3%	78.8%	3.5%	79.7%	2.8%
	Acc.	71.7%	7.0%	75.1%	2.5%	75.9%	1.6%	75.7%	1.6%
	SSD	7.1%	3.2%	6.3%	2.5%	5.8%	1.5%	5.9%	1.0%
BMC	Sens.	73.9%	10.5%	75.3%	6.2%	75.1%	6.0%	74.8%	6.8%
	Spec.	68.9%	7.0%	71.3%	6.0%	68.1%	8.1%	63.3%	8.6%
	Acc.	71.4%	7.1%	73.3%	3.3%	71.6%	3.9%	69.1%	4.3%
	SSD	7.9%	3.6%	7.8%	2.9%	8.8%	3.2%	10.0%	3.8%
Mean	Sens.	51.7%	11.5%	59.7%	6.1%	64.1%	4.5%	67.2%	3.9%
	Spec.	87.3%	2.4%	84.0%	1.9%	81.3%	1.6%	79.0%	1.3%
	Acc.	69.5%	5.0%	71.8%	2.8%	72.7%	2.1%	73.1%	1.7%
	SSD	18.1%	6.4%	13.5%	2.9%	11.6%	1.7%	10.8%	1.4%
Outlier	Sens.	64.9%	14.4%	74.0%	7.1%	78.2%	3.7%	80.3%	2.6%
	Spec.	71.7%	6.9%	64.0%	5.2%	58.5%	4.0%	54.3%	3.1%
	Acc.	68.3%	4.9%	69.0%	3.1%	68.3%	2.4%	67.3%	2.0%
	SSD	13.4%	5.8%	11.4%	2.1%	12.2%	1.7%	13.9%	1.4%
Combined	Sens.	66.6%	13.4%	74.9%	6.4%	78.6%	3.9%	80.6%	2.9%
	Spec.	68.1%	7.3%	60.1%	5.2%	54.3%	3.8%	49.9%	2.8%
	Acc.	67.4%	4.4%	67.5%	3.0%	66.5%	2.4%	65.3%	1.8%
	SSD	12.8%	5.2%	12.0%	1.9%	13.7%	2.0%	15.9%	1.8%

Sensitivity increased and specificity decreased for each conventional method (Means, Outlier, and Combined) as the number of features used for categorization increased. Sensitivity and specificity increased or remained steady for the AMC Probabilistic method as the number of features used for categorization increased.

ratio (AAR), otherwise known as thickness,[19] and revised size index (SI)[20] were also used as features. All possible combinations of feature sets taken two, three, four, and five at a time were examined for sensitivity, specificity, and accuracy to determine which combination of features used simultaneously is the best choice for a sensitive and specific muscle categorization. Each different feature set was subjected to 10 trials, where the set of MUP training data used per trial was the same from feature set to feature set to ensure consistent comparison across feature sets. The sensitivity across patient muscles and the specificity across control muscles were recorded for each trial. Accuracy and SSD were then calculated for each trial.

RESULTS

Tables 1 and 2 show the mean sensitivity, specificity, accuracy, and SSD across all possible feature sets comprised of two, three, four, or five features for the biceps brachii and FDI muscles, respectively.

When the results across all feature sets of five for the biceps brachii muscles were compared (the last column of Table 1) the AMC Probabilistic method had the highest mean accuracy of 75.7%, which differed from the Means and BMC method, whose mean accuracies were 73.1% and 69.1%, respectively ($P < 0.01$). When the results across all feature sets of five for the FDI muscles were compared (the last column of Table 2) the BMC Probabilistic method had the highest mean accuracy of 88.0%, which was not significantly different than the AMC method, with an accuracy of 87.7% ($P > 0.40$). It was significantly different from the Means method, which had an accuracy of 80.0% ($P < 0.01$). The Outlier and Combined methods had lower mean accuracy than the Means and Probabilistic methods regardless of the number of features in a feature set for both the biceps brachii and FDI muscles as shown in Tables 1 and 2.

As the number of features used for categorization increased, sensitivity increased while specificity decreased for the conventional methods. For example, the Outlier method sensitivity increased from 64.9% to 80.3%, and specificity decreased from 71.7% to 54.3% for two and five features per set, respectively, in Table 1. For the biceps brachii muscles, the conventional methods had a greater difference between sensitivity and specificity as highlighted by their large SSD values in Table 1 as compared to the AMC method. With the Probabilistic methods, sensitivity and specificity demonstrated greater consistency, unlike the conventional

Table 2. FDI: average performance across all feature sets.

		# Features used for testing							
		Two		Three		Four		Five	
Method		Mean	SD	Mean	SD	Mean	SD	Mean	SD
AMC	Sens.	79.6%	8.0%	81.4%	3.3%	81.6%	2.9%	81.0%	3.1%
	Spec.	88.5%	9.7%	92.4%	2.9%	93.6%	1.8%	94.4%	1.5%
	Acc.	84.1%	8.6%	86.9%	2.7%	87.6%	2.1%	87.7%	2.1%
	SSD	4.8%	2.3%	5.7%	1.6%	6.1%	1.2%	6.7%	1.3%
BMC	Sens.	79.5%	7.9%	81.6%	4.1%	83.3%	4.1%	84.5%	3.7%
	Spec.	88.0%	9.5%	91.0%	3.4%	91.0%	3.4%	91.6%	3.7%
	Acc.	83.8%	8.4%	86.3%	3.1%	87.1%	3.2%	88.0%	3.1%
	SSD	4.6%	2.5%	5.2%	2.0%	5.0%	1.6%	5.2%	1.2%
Mean	Sens.	74.2%	12.3%	79.1%	3.1%	80.3%	1.5%	80.8%	0.8%
	Spec.	89.0%	3.3%	85.0%	2.9%	81.8%	2.5%	79.1%	2.0%
	Acc.	81.6%	5.1%	82.1%	1.4%	81.0%	1.1%	80.0%	0.9%
	SSD	8.9%	6.7%	6.1%	1.4%	5.8%	0.7%	6.1%	0.6%
Outlier	Sens.	59.4%	13.7%	65.6%	9.4%	69.0%	6.0%	70.9%	3.1%
	Spec.	78.3%	3.8%	71.3%	3.9%	65.7%	4.0%	61.1%	3.8%
	Acc.	68.9%	6.7%	68.4%	5.4%	67.4%	4.1%	66.0%	3.0%
	SSD	10.3%	6.7%	5.8%	3.3%	4.9%	1.3%	5.8%	1.0%
Combined	Sens.	75.9%	9.3%	80.0%	3.0%	81.3%	1.7%	81.9%	1.3%
	Spec.	71.4%	5.0%	62.5%	4.4%	55.7%	3.9%	50.5%	3.2%
	Acc.	73.6%	4.1%	71.2%	2.4%	68.5%	2.1%	66.2%	1.8%
	SSD	6.3%	3.5%	9.0%	2.5%	12.8%	2.1%	15.7%	1.7%

Sensitivity increased and specificity decreased for each conventional method (Means, Outlier, and Combined) as the number of features used for categorization increased. Sensitivity and specificity increased or remained steady for the AMC Probabilistic method as the number of features used for categorization increased.

methods, as the number of features used for categorization increased.

Table 3 shows the best five feature sets as sorted by accuracy for the biceps brachii muscles. The BMC method had the feature set with the highest accuracy (79.7%), which was not significantly different from the AMC method [accuracy of 78.7% ($P > 0.2$)] and the Means method [accuracy of 76.4% ($P > 0.05$)]. The AMC method had a feature set with an SSD value of 1.7%, which was among the lowest SSD values of all of the top five feature sets of the methods.

Table 4 shows the best five feature sets as sorted by accuracy for the FDI muscles. The BMC method had the feature set with the highest accuracy of 94.6%, which was significantly different than the AMC method [accuracy of 90.9% ($P < 0.01$)] and the Means method [accuracy of 85.8% ($P < 0.01$)]. Also, the Probabilistic methods appear to be more robust against variations in the training data as shown by the generally lower SD of accuracy for the top five feature sets.

All of the feature sets were sorted according to accuracy. A paired *t*-test was conducted between the most accurate feature set and each subsequent feature set in the ordered list starting with the second most accurate, to find the feature sets whose accuracies were similar (i.e., $P > 0.05$). In Tables 5 and 6 "# feat sets" refers to the number of feature sets per method that had similar accuracy as described above. Tables 5 and 6 show the number of occurrences of each individual feature in the feature sets whose accuracy did not differ significantly ($P > 0.05$). Table 5 shows that the AMC method favored thickness, SI, phases, and turns as the most discriminative features, while the BMC method favored phases and thickness. The Means method favored turns, area, and thickness based on the number occurrences in the feature sets whose accuracy did not vary significantly for the biceps brachii muscles. Table 6 shows that the AMC method favored SI and duration as the most discriminative features, while the BMC method favored duration, area, and SI. The Means method favored SI and phases based on the number occurrences in the feature sets whose accuracy did not vary significantly for the FDI muscles.

DISCUSSION

During a routine needle EMG examination, detecting a single MUP with clearly neuropathic characteristics is not sufficient to confidently conclude that the muscle has been affected by a neuropathic

Table 3. Biceps brachii: best five feature sets per method.

Method	Feature sets	Sens (%)	Sens SD	Spec (%)	Spec SD	Acc (%)	Acc SD	SSD	SSD SD
AMC	thick phases turns	77.8%	0.0%	79.7%	3.8%	78.7%	1.9%	1.7%	1.2%
	area phases	83.3%	7.9%	73.8%	6.9%	78.6%	3.2%	7.0%	3.9%
	thick SI phases turns	77.8%	0.0%	79.3%	7.8%	78.5%	3.9%	3.4%	1.6%
	dur thick SI phases	77.8%	7.4%	79.3%	7.6%	78.5%	3.9%	5.3%	3.2%
	amp thick phases turns	77.8%	0.0%	79.3%	5.6%	78.5%	2.8%	2.2%	1.8%
BMC	amp thick phases turns	82.2%	5.7%	77.2%	6.7%	79.7%	3.5%	4.7%	3.1%
	amp thick phases	82.2%	5.7%	77.2%	7.1%	79.7%	3.3%	4.8%	3.4%
	area phases	84.4%	7.8%	73.1%	7.2%	78.8%	2.9%	7.9%	3.7%
	area phases turns	82.2%	7.8%	74.5%	6.5%	78.4%	3.8%	6.1%	3.5%
	thick phases turns	77.8%	0.0%	78.3%	4.0%	78.0%	2.0%	1.5%	1.2%
Mean	area turns	65.6%	16.1%	87.2%	8.5%	76.4%	5.3%	13.1%	8.8%
	area phases turns	65.6%	16.1%	86.9%	8.6%	76.2%	4.9%	13.3%	8.5%
	area thick turns	70.0%	11.8%	82.4%	7.7%	76.2%	5.2%	9.2%	4.7%
	thick turns	65.6%	8.2%	86.6%	5.7%	76.1%	4.6%	10.5%	5.4%
	area thick phases turns	70.0%	11.8%	82.1%	7.9%	76.0%	4.9%	9.3%	4.5%
Outlier	area phases turns	80.0%	13.7%	70.7%	16.8%	75.3%	7.7%	11.2%	7.7%
	area turns	80.0%	13.7%	70.7%	16.8%	75.3%	7.7%	11.2%	7.7%
	thick phases turns	71.1%	16.7%	77.6%	11.1%	74.3%	5.7%	12.1%	4.4%
	area phases	74.4%	10.5%	74.1%	14.7%	74.3%	6.3%	8.9%	5.9%
	thick turns	68.9%	19.5%	77.6%	11.1%	73.2%	6.8%	13.2%	5.7%
Combined	area phases turns	81.1%	12.9%	66.6%	17.8%	73.8%	8.1%	11.5%	9.4%
	area turns	81.1%	12.9%	66.6%	17.8%	73.8%	8.1%	11.5%	9.4%
	thick turns	74.4%	14.9%	73.1%	10.5%	73.8%	6.1%	9.2%	5.9%
	thick phases turns	74.4%	14.9%	72.8%	11.1%	73.6%	6.3%	9.3%	6.0%
	area phases	74.4%	10.5%	71.7%	14.8%	73.1%	6.0%	9.4%	5.7%

The top five feature sets per method as sorted by accuracy are presented.

Table 4. FDI: best five feature sets per method.

Method	Feature sets	Sens (%)	Sens SD	Spec (%)	Spec SD	Acc (%)	Acc SD	SSD	SSD SD
AMC	amp dur area SI phases	84.6%	0.0%	97.2%	4.2%	90.9%	2.1%	6.3%	2.1%
	amp dur area SI turns	83.8%	2.4%	97.6%	2.8%	90.7%	2.0%	6.9%	1.7%
	dur area SI phases turns	84.6%	0.0%	96.0%	1.9%	90.3%	0.9%	5.7%	0.9%
	amp dur SI phases turns	84.6%	0.0%	96.0%	3.3%	90.3%	1.6%	5.7%	1.6%
	dur area SI phases	84.6%	0.0%	96.0%	1.9%	90.3%	0.9%	5.7%	0.9%
BMC	dur area SI turns	90.8%	3.2%	98.4%	2.8%	94.6%	1.8%	3.8%	2.4%
	dur area SI	90.8%	3.2%	98.4%	2.8%	94.6%	1.8%	3.8%	2.4%
	amp dur area SI	87.7%	6.5%	98.8%	2.7%	93.2%	3.0%	5.6%	3.9%
	amp dur area SI turns	86.9%	6.3%	98.8%	2.7%	92.9%	2.9%	6.0%	3.9%
	dur area thick SI turns	89.2%	5.4%	96.4%	4.0%	92.8%	2.2%	4.0%	3.7%
Mean	SI phases	78.5%	3.2%	93.2%	6.0%	85.8%	2.7%	7.4%	4.0%
	area phases	80.0%	4.0%	90.0%	6.6%	85.0%	2.5%	6.3%	2.6%
	SI turns	78.5%	3.2%	91.2%	5.3%	84.8%	2.6%	6.4%	3.5%
	SI phases turns	78.5%	3.2%	90.8%	6.3%	84.6%	3.2%	6.3%	3.7%
	amp phases	76.9%	8.9%	91.6%	6.9%	84.3%	3.8%	8.3%	5.7%
Outlier	area SI	72.3%	4.0%	85.2%	9.4%	78.8%	3.7%	8.1%	3.3%
	area phases	72.3%	4.0%	81.6%	9.1%	77.0%	3.4%	6.4%	3.9%
	area SI phases	72.3%	4.0%	80.4%	10.4%	76.4%	3.9%	6.4%	4.4%
	area turns	72.3%	4.0%	79.2%	5.9%	75.8%	2.3%	4.8%	2.7%
	dur area	72.3%	4.0%	79.2%	7.5%	75.8%	3.8%	4.7%	3.2%
Combined	area SI	81.5%	4.0%	77.6%	9.5%	79.6%	4.3%	4.0%	4.5%
	area phases	81.5%	4.0%	75.2%	9.6%	78.4%	4.5%	4.5%	4.7%
	SI phases	78.5%	3.2%	77.2%	8.4%	77.8%	3.3%	3.9%	3.7%
	area SI phases	81.5%	4.0%	72.8%	10.3%	77.2%	4.6%	5.4%	5.3%
	amp area	83.1%	3.2%	71.2%	7.7%	77.1%	3.2%	6.6%	4.1%

The top five feature sets per method as sorted by accuracy are presented.

Table 5. Biceps: number of occurrences per feature in best feature sets.

	AMC	BMC	MEAN	Outlier	Combined
amp	8	3	21	10	10
dur	11	1	18	5	2
area	7	4	32	22	17
thick	17	6	30	15	11
SI	14	0	18	15	9
phases	14	8	22	19	16
turns	13	4	41	24	18
# feat sets	23	9	48	36	29

The number of occurrences of each individual feature in the feature sets whose accuracy did not differ significantly from the most accurate feature set is shown per each method. The term "# feat sets" refers to the number of feature sets per method that had the same accuracy at a 5% level of significance.

process. This is because MUPs with neuropathic feature values (e.g., long duration, large area, large thickness) can be found in normal or abnormal muscles, but with different probabilities. As such, single MUPs in isolation do not provide enough information about the pathophysiological state of a muscle to support a clinical decision. However, combining information provided by a set of MUP characterizations can provide adequate information for a representative muscle characterization to support an objective clinical diagnosis—a process that is inherent with Probabilistic muscle characterization.

Data for the FDI muscle from a patient with a neuropathic disorder, selected from the data described in Materials and Methods, was used to demonstrate the Probabilistic muscle characterization methods (see Fig. 1). Along the left edge of Figure 1 the template waveform and characterization of each of the 13 MUPs detected from the FDI muscle are shown. Each MUP characterization is represented by a pie chart where the blue area is proportional to the conditional probability that the MUP was detected in a neuropathic muscle and the green area is proportional to the conditional probability that the MUP was detected in a normal muscle. The 13 MUPs were sorted in order from highest conditional probability of being neuropathic to lowest. For the AMC method the MUP characterization values per category were averaged (see eq. 1), resulting in a muscle characterization measure of 82% in support of a neuropathic condition and 18% in support of a normal condition. Using eq. 2, the BMC method calculated a muscle characterization measure of 100% in support of a neuropathic condition and 0% in support of a normal condition. In both cases this muscle is correctly categorized as neuropathic. MUP 13 is useful for establishing relative scale, because it is the sole MUP in the set whose probability of being detected in a normal muscle was higher than in a neuropathic muscle. MUPs 1–12 in Figure 1 have relatively large amplitudes compared to MUP 13.

Probabilistic muscle characterization resembles the methods clinicians use to qualitatively examine needle EMG signals. First, a clinician subjectively assesses the similarity of a specific MUP under examination to MUPs detected in muscles of specific categories and then implicitly forms an estimate of the probability of detecting this MUP in a muscle of a specific category. This is similar to the MUP characterizations represented by the smaller pie charts in Figure 1. Next, the clinician combines the probability estimates of all of the MUPs examined to formulate an overall muscle characterization—similar to the muscle characterization represented by the larger pie chart to the right in Figure 1. However, clinicians may be prone to making biased decisions, because they may look for MUPs that confirm a preconceived expectation or assign lesser importance to MUPs that contradict an expectation. As well, there is a possibility of quickly jumping to an incorrect conclusion based on the observation of a single MUP feature, (e.g., it is common to associate increased MUP amplitude with a neuropathic condition). Probabilistic muscle characterization is a quantitative method that uses unbiased MUP characterizations that are estimated by simultaneously considering multiple MUP features and based on numbers of occurrences in exemplary training data. A summary of the differences between clinicians using qualitative analysis and Probabilistic methods is provided in Table 7.

With the Probabilistic methods further details relating to the rationale of each MUP characterization are available if desired. Appendix B shows the waveform, feature values, subsets of feature values

Table 6. FDI: number of occurrences per feature in best feature sets.

	AMC	BMC	MEAN	Outlier	Combined
amp	7	2	3	0	1
dur	12	4	0	0	0
area	9	4	5	3	4
thick	2	0	2	0	0
SI	13	4	7	2	3
phases	5	0	7	2	3
turns	7	2	2	0	0
# feat sets	13	4	11	3	5

The number of occurrences of each individual feature in the feature sets whose accuracy did not differ significantly from the most accurate feature set is shown per each method. The term "# feat sets" refers to the number of feature sets per method that had the same accuracy at a 5% level of significance.

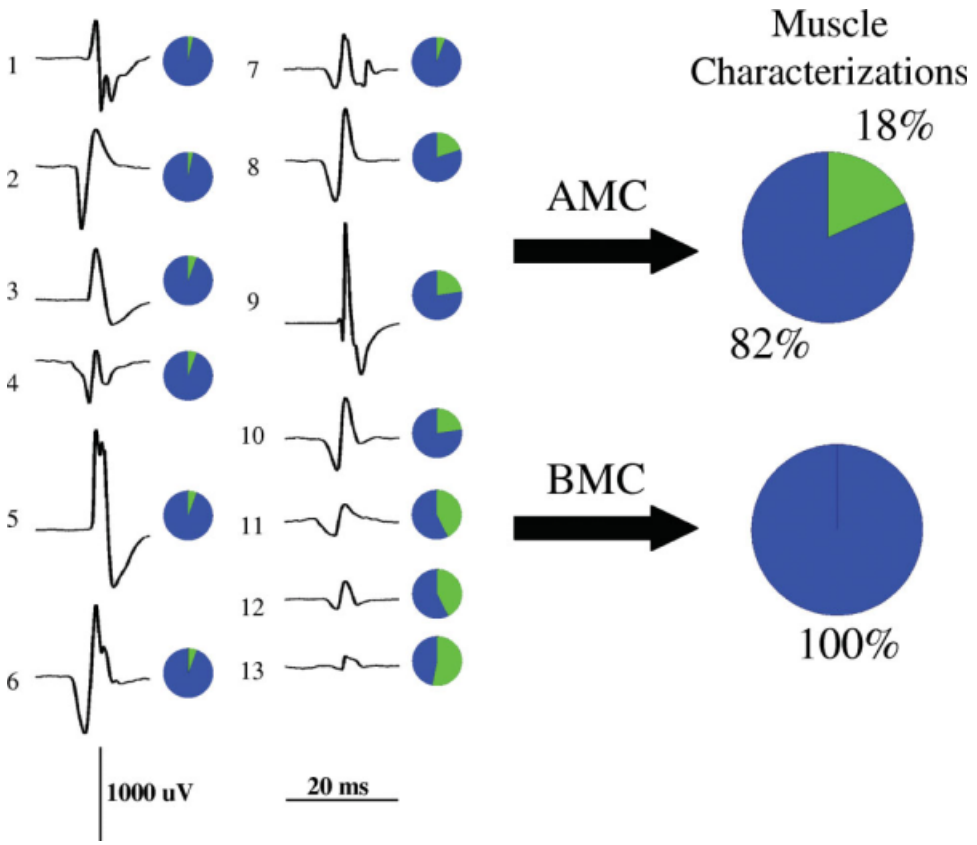


FIGURE 1. Example of FDI muscle characterization using Probabilistic methods. Data from an FDI muscle whose actual diagnosis is neuropathic was used to demonstrate the Probabilistic muscle characterization methods. The MUP characterizations in this figure were based on duration, area, and thickness feature values. MUP feature values were measured using a 50-ms sweep. They are truncated visually in the figure to conserve space. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

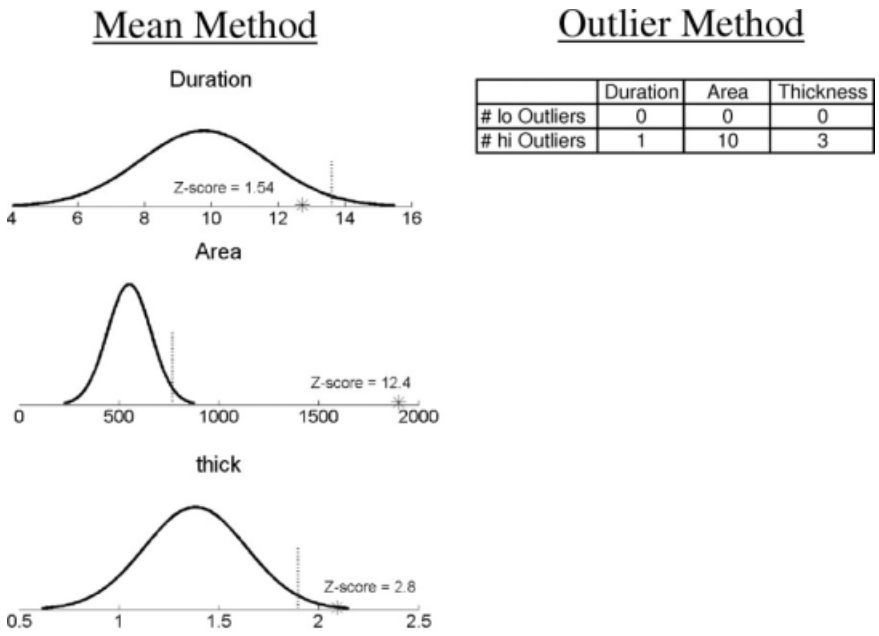
supporting normal, neuropathic, and the MUP characterization for the MUPs in Figure 1 that had the highest conditional probability supporting a neuropathic characterization and the two lowest conditional probabilities supporting a neuropathic characterization. Appendix B demonstrates the transparency of the PD method for MUP characterization,[10] i.e., the justification for an MUP characterization is given by the set of MUP feature values that support or refute each category as well as the strength of support or refutation. Aside from the actual use of a quantitative muscle characterization, simultaneously considering multiple MUP features and estimating specific muscle category occurrence probabilities increases the cognitive burden for the electromyographer during qualitative examination of MUPs. As such, it is expected that the speed and accuracy with which electromyographers master the difficult skills involved with qualitative muscle characterization could be improved with these types of graphical explanations of quantitative MUP characterization.

PD used three intervals to quantize feature values in this work and performed well relative to the other methods examined. Using a small number of intervals helps to simplify the visual patterns

Table 7. Qualitative EMG examination compared to Probabilistic method.

Clinician	Probabilistic method
• qualitative • prone to bias • estimations based on examiner experience	• quantitative • objective • MUP characterization estimation based on number of occurrences in exemplary training data
• large cognitive effort needed to simultaneously consider multiple feature values across many MUPs	• computer automated

Figure 1 can also be used as an analogy of how clinicians qualitatively examine EMG signals. How a clinician performs a qualitative electromyographic examination compared to the Probabilistic method is shown in this table.



	Duration	Area	Thickness
# lo Outliers	0	0	0
# hi Outliers	1	10	3

FIGURE 2. Example of Means and Outlier analysis of an FDI Muscle. The Means and Outlier methods are performed using the same muscle as in Figure 1. The curves for the Means method were drawn assuming a Gaussian distribution of the feature values of the MUPs detected from normal muscles. The mean and SD parameters used for drawing the curves were calculated using the mean and SD of feature values as described in Means and Outlier Methods in Materials and Methods. The curves are meant to provide an indication of the variance of each feature value and do not represent their actual distributions because the feature values are not exactly Gaussian distributed, but are assumed to be Gaussian for the Means analysis. The dashed vertical line to the right of the mean of each curve is the mean plus two SD, i.e., the upper threshold of the normative training data. The asterisk in each plot is where the mean for each feature value for the muscle under examination fell. The mean value of duration across MUPs detected from the test muscle is less than 2 SD, so it falls in the normative range. The means of area and thickness of the test muscle are 12.4 and 2.8 SD higher than the mean value of the normative training data for each feature, respectively, indicating a neuropathic condition. The table headed Outlier Method shows that area and thickness each have three or greater outliers, which are both indications of neuropathic muscle. As such, both methods correctly categorize this muscle as neuropathic.

that explain the results leading to diagrams that are easily recognized and understood (see Appendix B). Using a small number of intervals results in a wider range of feature values that are considered small, medium, or large. This suggests that for Probabilistic muscle characterization, high levels of precision in placing the onset and end markers of MUPs are not required to ultimately achieve a high level of muscle categorization accuracy.

Using the same muscle as in Figure 1, Figure 2 demonstrates the Means and Outlier methods. Figure 2 shows that mean values of area and thickness are greater than 2 SD above the normative training data, indicating a neuropathic condition. The table headed Outlier Method in Figure 2 shows that area and thickness each have three or greater outliers, which are both indications of neuropathic muscle. As such, both methods correctly categorize this muscle as neuropathic. This study extends previous work using external anal sphincter (EAS) muscle data,[6] in that data from two additional muscles are examined that are affected by two disease processes. Additionally, it uses DQEMG versus Multi-MUP decomposition techniques to obtain quantitative MUP data, and it evaluates a new method of combining MUP characterizations (AMC) to obtain a muscle characterization.

Overall, the data sampled for this work resulted in higher characterization accuracy for FDI muscles compared to biceps brachii muscles (as shown in Tables 3 and 4). This general result is consistent with the tendency of neuropathic conditions including ALS to generally affect distal muscles to a greater extent than proximal muscles. A higher degree of involvement corresponds to greater accuracy in characterizing a muscle because of the greater separation between the distributions of control and patient data values. In this study, the Probabilistic methods performed better than the conventional methods. With the EAS data the BMC method performed slightly better than the conventional methods.[6] (The AMC method was not analyzed previously.[6]) Although in this study the accuracy of the Means method was

almost as high as the Probabilistic methods, the Probabilistic methods had lower SD values. Furthermore, with the EAS data, the Means method had very poor sensitivity, i.e., 39% average sensitivity across all sets of four features, as discussed previously.[6] In this study and with the EAS data, sensitivity and specificity remained high with lower SSD and SD values for the Probabilistic methods regardless of the number of features used for characterization. The performance of a method should be evaluated across different clinical data and muscle types. The separation between the feature value distributions of normal and abnormal MUPs for the EAS data[6,7] was less than the separation between the distributions for these data. The Probabilistic method performed as well or better than any of the conventional methods with the EAS data. However, the conventional methods were not consistent across different datasets. The performance of the Combined method with the EAS data was comparable to that of the BMC Probabilistic method. However, with these data the Combined method performed more poorly than the Probabilistic method. As well, as depicted in Figure 2, the Means and Outlier methods provide only a dichotomous indication as to whether a muscle is normal or abnormal, complicating their use for determining the level of involvement of a disorder or longitudinal studies. Alternatively, plots such as those presented in Figure 1 are more easily interpreted with regard to confidence and level of involvement and suggest that the Probabilistic methods can be much more easily used to support "possible," "probable," or "definite" determinations of levels of pathology and to interpret changes across longitudinal studies.

In previous work, when the Means and Outlier methods were used together, testing of a muscle required a minimum of 20 MUPs.[6,7] In this work, all of the muscles were tested regardless of the number of MUPs detected per muscle. Only one biceps brachii muscle from the control subjects had less than 20 MUPs. The biceps brachii data had 1 out of 9 and the FDI data had 2 out of 13 muscles from patients with less than 20 MUPs. Allowing all muscles to be tested does not require that the clinician keep looking until a minimum of 20 MUPs are found—a potentially difficult task in muscles where a neuropathic process reduces the number of motor units. Note that the Probabilistic method requires a smaller number of MUPs to be detected than the conventional methods to test a muscle.

Using training data from both patient and control muscles gives the Probabilistic methods additional information compared to the conventional methods that use only control data for training. This additional information provides the Probabilistic methods with the ability to report a characterization measure (similar to probability) for each category and to achieve slightly higher accuracy and better balance between specificity and sensitivity as compared to the conventional methods. The Probabilistic methods can provide a single characterization measure for each category from a range of continuous values and renders them potentially useful for determining the level of involvement of a disorder. The characterization measures of the AMC method showed excellent correlation with level of involvement of neuropathic and myopathic neuromuscular disorders when simulated electromyographic signals were analyzed based on a biological model.[21] The Results section and previous work[5,6] also show that the BMC method is an accurate method; however, the muscle characterization measures produced often saturate at either 0% or 100%. This saturation is a characteristic of this particular method and has been previously discussed.[18]

The Probabilistic characterization method is easily extended to handle additional categories, i.e., myopathic disorders. Results when adding a myopathic category are currently being compiled and will be presented in future work. As well, other QEMG measures can be included in future works as sources of additional information for determining an overall muscle characterization. And finally, although not a focus of this study, the results of the analysis of a set of individual muscles may be combined to provide a characterization for a patient or a given neuromuscular system.

In conclusion, this work suggests that the Probabilistic methods are well suited to form the basis of a neuromuscular clinical decision support system, as the results showed higher accuracy and lower SSD than the conventional Means and Outlier methods. This leads to the expectation that clinicians who use the Probabilistic method will have better clinical decision support.

In addition, the Probabilistic method can facilitate the determination of "possible," "probable," or "definite" levels of pathology, whereas the conventional Means and Outlier methods can only provide a dichotomous "normal" or "abnormal" decision.

APPENDIX A

Pattern Discovery-Based MUP Characterization. An MUP can be characterized using rules discovered by PD applied to a set of MUP training data.[22–25] Initially, each

MUP is described by its continuous feature values, (e.g., amplitude, duration, number of phases, etc.). The ranges of each feature's values are divided into a number of predefined intervals (in this work three intervals were used), and each interval is given a natural language label such as small, medium, or large. Each feature value is then assigned a discrete value (i.e., quantized) based on the interval in which its continuous value falls. Given the training dataset with hundreds of MUP samples from each muscle category, for each category the number of occurrences of all subsets of quantized feature values across the number of MUP samples are counted. If the number of occurrences of a subset is significantly higher or lower than the number of occurrences expected assuming randomly distributed data, then the subset of quantized feature values is said to be a component rule for that category.

The i^{th} component rule for a category is denoted by $R_i(category)$, and its discriminatory power can be determined by its weight of evidence (WOE), which is the odds of a sample MUP having a subset of quantized features values matching the set of quantized feature values of $R_i(category)$ belonging to that category versus not belonging to that category. The following equations provide the WOE for normal and neuropathic, respectively:

$$WOE(norm) = \log_2 \frac{P(R_i(norm))}{P(R_i(neur))} \quad (A.1)$$

$$WOE(neur) = \log_2 \frac{P(R_i(neur))}{P(R_i(norm))} \quad (A.2)$$

Note that the range of eqs. A.1 and A.2 is $-\infty < WOE < +\infty$. A positive value for $WOE(norm)$ indicates that the component rule provides support for the categorization of normal. A negative value for $WOE(norm)$ indicates that the component rule refutes the categorization of normal. The strength of the support (or refutation) is proportional to the absolute value of the WOE. In many cases, component rules do not include all of the features available for characterization. In those cases, the WOE is determined by summing the WOE of the set of component rules that include only component rules that do not duplicate any features for a category.

The WOE for a category can be converted into a conditional probability using the following equation:

$$PMUP(category) = \frac{1}{\left[(2^{-WOE}) \cdot \left(\frac{1-P_0(category)}{P_0(category)}\right)\right] + 1} \quad (A.3)$$

where category = muscle category (normal or neuropathic); $PMUP(category)$ = the MUP characterization value for that *category* of the MUP detected from the muscle under test (i.e., P(*category* | MUP)); and $P_0(category)$ = the prior probability of the category.

An example will help clarify. Say that the features duration, area, and thickness are used for characterizing an MUP using two categories: normal or neuropathic. In this example, each of the feature values are quantized into one of three intervals called low, medium, or high. Also, a prior probability of 0.5 is assumed for each category. Consider MUP 1 in Figure 3 that has medium duration, high area, and high thickness. The feature values are compared against the component rules that were discovered in the training data, and two component rules are found—one for the normal category and the other for the neuropathic category. The $WOE(neur)$ was calculated by counting 29 occurrences in the training data of the neuropathic MUPs that these feature values appear and counting 1 MUP in the normal training data with those same feature values to empirically determine the probabilities of belonging to the normal and neuropathic categories. There were 414 MUPs in the normal training data and 414 MUPs in the neuropathic training data, and so using eqs. A.1 and A.2:

$$WOE(norm) = \log_2 \frac{P(R_i(norm))}{P(R_i(neur))} = \log_2 \frac{1/414}{29/414} = -4.86$$

$$WOE(neur) = \log_2 \frac{P(R_i(neur))}{P(R_i(norm))} = \log_2 \frac{29/414}{1/414} = +4.86$$

Using eq. A.3 converts the WOEs into MUP conditional probabilities of $PMUP(norm) = 0.03$, and $PMUP(neur) = 0.97$, respectively. Further information on PD-based MUP characterization can be found in previous work.[6,8–10]

APPENDIX B

Detailed MUP Characterizations: Probabilistic Method (FDI Muscle). Figure 3 demonstrates the transparency of the PD method for MUP characterization by showing the set of MUP feature values that support or refute each category as well as the strength of support or refutation.[10] The features duration, area, and thickness were used in this example, because this feature set was among the set of feature sets whose accuracy did not differ significantly from the best accuracy; it had only three features, which simplifies visual interpretation of the MUP characterizations; it was among the most robust, with a low SD of accuracy of 1.6% across 10 trials; and it had low SSD, with a mean value of 2.8% across all 10 trials. The second column shows the value of the feature prior to quantization (assignment of a continuous value to a discrete value based on the interval in which its continuous value falls; see Appendix A for further explanation) and how each of the feature values was quantized into an interval of low, medium, or high as indicated by the length of each bar. The third and fourth columns show the sets of feature values that support or refute normal and neuropathic, respectively. A

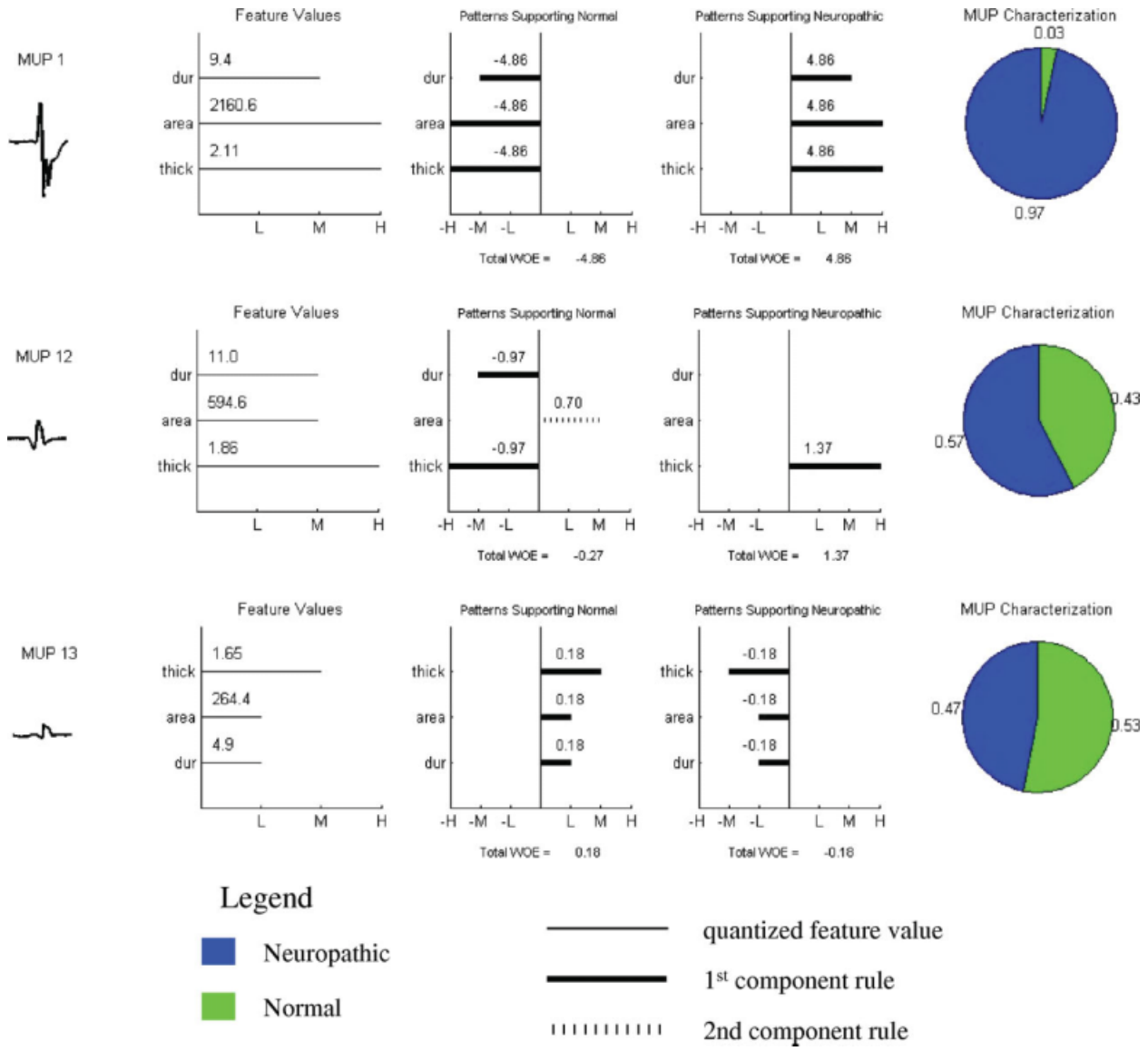


FIGURE 3. Detailed MUP characterizations using the Probabilistic method (FDI muscle). Each row in the figure is an example of the detailed information that the PD method can provide to explain its characterization of MUPs. The second column shows the feature values and their corresponding quantized interval of either low (L), medium (M), or high (H). The set of features that together support or refute a category is known as a component rule, i.e., pattern. The third and fourth columns show the set of component rules used to characterize an MUP. Different component rules are shown using different line styles as drawn in the legend. The total WOE is calculated by taking the sum of the component rules for a category. The total WOE for a category is transformed into a conditional probability shown in the pie chart at the right using eq. A.3. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

set of bars with the same line style pointing to the right show a set of feature values that support a category, while a set of bars with the same line style pointing to the left show a set of feature values refuting a category. A set of features that together support or refute a category is called a pattern, i.e., component rule. As an example, a clinician could use patterns to understand why MUP 13 has almost equal conditional probabilities of being detected from a normal and neuropathic muscle of 0.53 and 0.47, respectively. The patterns that support normal for MUP 13 show that low duration, low area, and medium thickness, support normal with a marginal WOE of 0.18. The WOE of 0.18 means that the odds of finding an MUP with those feature values in the normal training data are $2^{\wedge 0.18} = 1.09$ times greater than finding an MUP with those feature values in the neuropathic training data (see Appendix A for further details on how WOE is calculated). The column titled "patterns supporting neuropathic" shows a WOE of -0.18 that refutes neuropathic categorization based on the odds that finding an MUP with those feature values in the neuropathic training data are 0.92 times less than finding an MUP with those feature values in the normal training data.

Because the strength of support for normal based on a WOE of 0.18 did not have a large difference compared to the WOE refuting neuropathic the conditional probabilities for normal and neuropathic are almost equal in value (0.53 and 0.47, respectively).

REFERENCES

1. Kendall R, Werner RA. Interrater reliability of the needle examination in lumbosacral radiculopathy. Muscle Nerve 2006;34:238–241.
2. Swash M. What does the neurologist expect from clinical neurophysiology? Muscle Nerve 2002;999:S134.
3. Stålberg E, Bischoff C, Falck B. Outliers, A way to detect abnormality in quantiative EMG. Muscle and Nerve 1994;17: 392–399.
4. Stewart CR, Nandedkar SD, Massey JM, Gilchrist JM, Barkhaus PE, Sanders DB. Evaluation of an automatic method of measuring features of motor unit action potentials. Muscle Nerve 1989;12:141–148.
5. Pfeiffer G. The diagnostic power of motor unit potential analysis: an objective Bayesian approach. Muscle Nerve 1999;22:584–591.
6. Pino LJ, Stashuk DW, Podnar S. Bayesian characterization of external anal sphincter muscles using quantitative electromyography. Clin Neurophysiol 2008;119:2266–2273.
7. Podnar S. Criteria for neuropathic abnormality in quantitative anal sphincter electromyography. Muscle Nerve 2004;30:596.
8. Hamilton-Wright A, Stashuk DW, Pino LJ. On weight of evidence based reliability in pattern discovery. In: Proceedings of the 11th International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems IPMU, IEEE/IFSA 2006.
9. Hamilton-Wright A, Stashuk DW, Pino LJ. Internal measures of reliability in 'pattern discovery' based fuzzy inference. In: Proceedings of the 11th International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems IPMU 2006.
10. Pino LJ, Stashuk DW, Boe SG, Doherty TJ. Motor unit potential characterization using "pattern discovery." Med Eng Phys 2008;30:563–573.
11. Brooks BR, Miller RG, Swash M, Munsat TL. El Escorial revisited: revised criteria for the diagnosis of amyotrophic lateral sclerosis. Amyotroph Lateral Scler Other Motor Neuron Disord 2000;1:293–299.
12. 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;31:365–373.
13. Boe SG, Stashuk DW, Doherty TJ. Motor unit number estimates and quantitative motor unit analysis in healthy subjects and patients with amyotrophic lateral sclerosis. Muscle Nerve 2007;36:62–70.
14. 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;29:693–699.
15. Boe SG, Stashuk DW, Doherty TJ. Within-subject reliability of motor unit number estimates and quantitative motor unit analysis in a distal and proximal upper limb muscle. Clin Neurophysiol 2006;117:596–603.
16. Stashuk DW. Decomposition and quantitative analysis of clinical electromyographic signals. Med Eng Phys 1999;21: 389–404.
17. Stålberg E, Falck B, Sonoo M, Stålberg S, Åström M. Multi-MUP EMG analysis — a two year experience in daily clinical work. Electroencephalogr Clin Neurophysiol Electromyogr Motor Control 1995;97:145–154.
18. Duda R, Hart PE. Pattern classification, 2nd ed. Hoboken, NJ: John Wiley & Sons; 2001.
19. Nandedkar SD, Barkhaus PE, Sanders DB, Stålberg EV. Analysis of amplitude and area of concentric needle EMG motor unit action potentials. Electroencephalogr Clin Neurophysiol 1988;69:561–567.
20. Sonoo M. New attempts to quantify concentric needle electromyography. Muscle Nerve 2002;999:S98–S102.
21. Pino LJ, Stashuk DW. Using motor unit potential characterizations to estimate neuromuscular disorder level of involvement. 30th Annual International IEEE EMBS Conference, Vancouver, British Columbia, Canada, August 2008.
22. Wang Y. High-order pattern discovery and analysis of discrete-valued data sets. University of Waterloo, PhD Thesis; 1997.
23. Wong AKC, Wang Y. Discovery of high order patterns. In: Proceedings of the 1995 IEEE International Conference on Systems, Man and Cybernetics 1995;2:1142–1147.
24. Wong AKC, Wang Y. High-order pattern discovery from discrete-valued data. IEEE Trans Knowled Data Eng 1997;9: 877–893.
25. Wong AKC, Wang Y. Pattern discovery: a data driven approach to decision support. IEEE Trans Syst Man Cybern C Appl Rev 2003;33:114–124.